Vol. 40, No. 4 August, 1950

Journal of Experimental Psychology

EDITED BY

FRANCIS W. IRWIN

UNIVERSITY OF PENNSYLVANIA

CONSULTING EDITORS

CLARENCE H. GRAHAM, *Columbia University*
HARRY HELSON, *Brooklyn College*
DONALD B. LINDSLEY, *Northwestern University*
DONALD G. MARQUIS, *University of Michigan*
GEORGE A. MILLER, *Harvard University*
L. L. THURSTONE, *University of Chicago*
G. R. WENDT, *University of Rochester*
JOSEPH ZUBIN, *N. Y. Psychiatric Institute, Columbia University*

CONTENTS

An Experimental Analogue of Repression: II. The Effect of Individual Failure and Success on Memory Measured by Relearning: A. F. ZELLER 411
The Effect of Different Types of Preliminary Activities on Subsequent Learning of Paired-Associate Material: L. E. THUNE 423
Stimulus Pre-Differentiation as a Factor in Transfer of Training: R. M. GAGNÉ AND KATHERINE E. BAKER 439
Retention as a Function of Stage of Practice: RUTH GREENBERG AND B. J. UNDERWOOD 452
A Quantitative Comparison of the Discriminative and Reinforcing Functions of a Stimulus: J. A. DINSMOOR 458
Resistance to Extinction of a Conditioned Operant as Related to Drive Level at Reinforcement: R. C. STRASSBURGER 473
The Effect of Extinction upon Rate of Reconditioning: J. DEESE 488
A Study of Concept Formation as a Function of Reinforcement and Stimulus Generalization: A. H. BUSS 494
The Learning and Retention of Concepts: V. The Influence of Form of Presentation: H. B. REED 504
Sensory Conditioning Measured by the Facilitation of Auditory Acuity: W. J. BROGDEN .. 512
The Relation of Vernier and Depth Discrimination to Width of Test Rod: R. N. BERRY, L. A. RIGGS AND W. RICHARDS 520
An Experiment on the Expressiveness of Shell and Textile 'Montages': IVY G. CAMPBELL-FISHER 523
An Hypothesis Concerning the Generation and Use of Synonyms: W. M. LEPLEY 527
The Relationship Between the Tilt of a Visual Field and the Deviation of Body Position from the Vertical in the White Rat: B. F. RIESS, H. KRATKA AND A. DINNERSTEIN .. 531

PUBLISHED BI-MONTHLY BY

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

PRINCE AND LEMON STS., LANCASTER, PA.

AND 1515 MASSACHUSETTS AVE., N. W., WASHINGTON 5, D. C.

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879
Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. & R. of 1948, authorized September 11, 1947.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published bi-monthly by the American Psychological Association, Inc. The annual subscription rate is $7.00 (foreign, $7.50); single copies, $1.25. Subscriptions, orders, and business communications should be addressed to the American Psychological Association, Inc., 1515 Massachusetts Ave., N.W., Washington 5, D. C.

This JOURNAL publishes articles which are intended to contribute toward the development of psychology as an experimental science. Experimental work with normal human subjects is favored over work with abnormal or animal subjects. Studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental psychology or for psychological theory. Articles of a length exceeding 25 printed pages cannot be accepted.

With rare exceptions, articles are published in the order of their receipt. Authors are supplied with 50 free offprints without covers. All of the cost of author's alterations in galley proof is charged to the author, and one-half of the cost of tables, cuts, and formulas.

Immediate publication can be given to articles whose authors assume the full cost of publication. Such articles, if submitted before October 1, 1950, should be addressed to the retiring editor, Francis W. Irwin, Department of Psychology, University of Pennsylvania, Philadelphia 4, Pa. No free reprints are given in the case of these articles.

Address articles submitted for regular publication, or for immediate publication after the December, 1950 issue, to the editor-elect, Arthur W. Melton, Human Resources Research Center, Lackland Air Force Base, San Antonio, Texas.

This JOURNAL adheres to the conventions concerning bibliographies, citations, preparation of tables and cuts, etc., as described by J. E. Anderson and W. L. Valentine, The preparation of articles for publication in the journals of the American Psychological Association, *Psychol. Bull.*, 1944, 41, 345–376. Special attention is called to the style of numbers and abbreviations (p. 353). The term 'experimenter,' with its various forms, singular, plural, and possessive singular and plural, is abbreviated *E*, *Es*, *E*'s, and *Es*'; analogous abbreviations are used for the terms 'subject' (*S*) and 'observer' (*O*). All manuscripts must be typed double-spaced throughout, *including tables, footnotes, and references.* Tables, figures, and footnotes must be placed on separate pages at the end of the manuscript.

Journal of
Experimental Psychology

VOL. 40, No. 4 AUGUST, 1950

AN EXPERIMENTAL ANALOGUE OF REPRESSION. II. THE EFFECT OF INDIVIDUAL FAILURE AND SUCCESS ON MEMORY MEASURED BY RELEARNING[1]

ANCHARD FREDERIC ZELLER[2]

The Johns Hopkins University

In a recent review article by Zeller (7), the following design was proposed as the necessary framework for an experimental test of the concept of repression proper or after-expulsion:

	Control		Experimental
I.	Step 1. Learning	=	Learning
	Step 2. Retention test	=	Retention test
	Time Interval		
II.	Step 3. Retention test	=	Retention test
	Step 4. Neutral task		Repression
	Step 5. Retention test	>	Retention test
	Time Interval		
III.	Step 6. Retention test	>	Retention test
	Step 7. Neutral task		Removal of repression
	Step 8. Retention test	=	Retention test

The present experiments have followed this design.

INDIVIDUAL EXPERIMENT I

Relearning 15 Semi-Nonsense Syllables

In order to fulfill the conditions outlined it was necessary first to find some kind of neutral material to be learned and for which the rate of learning could be quantified. Second, it was necessary to find some kind of interpolated material which could be manipulated by the experimenter so that it would have either neutral, success or traumatic value to the subject. The first condition presented very little difficulty as there is a great variety of relatively neutral verbal material available. The paired associates method was adopted, using as the specific material the control list of Flanagan (1) consisting of 15 paired three-letter semi-nonsense syllables.

The choice of the interpolated material presented more difficulty. It seemed, however, that the best that could be chosen was some type which could be presented in a situation which was so unstructured that the subject was entirely dependent on the examiner for the interpretation of his performance. Further, the material had to be of such a nature that any statement made by the examiner about the subject's success or failure could be readily accepted because apparently confirmed by the objective evidence. After several other methods had been tried and discarded, the medium chosen for the artificial in-

[1] The present paper is a portion of a dissertation presented to the Faculty of Philosophy of the School of Higher Studies, The Johns Hopkins University, in partial fulfillment of the requirements for the Ph.D. degree.

[2] Now at the University of Tulsa.

duction of success or failure was imitative tapping using Knox cubes. Knox cubes consist of a set of five one-in. wooden cubes. These were chosen because the task of reproducing a tapping pattern with these blocks appears simple and yet even the best subjects are seldom able to imitate a pattern of eight units. As very few individuals are acquainted with the standard performance level with Knox cubes, there is little chance that the examiner's interpretation would be questioned by the subject.

Material.—The material for this experiment consisted, first, of a set of 15 3 × 5 in. cards on each of which was typed in capitals the syllable pair to be associated; second, a set of stimulus cards on each of which was typed the first of the paired syllables which appeared on the first set; and third, a set of Knox cubes (i.e., five one-in. black cubes). Timing was done by a clock pendulum with a 2.1-sec. swing or with a Waltham 10-sec. sweep stop watch. The only other material required was paper and pencil.

Subjects.—The subjects for this experiment were 20 male and female college students who volunteered their services. Although it is not possible to give an accurate figure, at least 50 percent of these volunteered in the hope of eventually obtaining psychological advice. It is therefore probable that they represent a more emotionally unstable group than a random selection of college students.

Procedure.—Each subject was tested individually and privately. In order that the true nature of the experiment might not be revealed to the subject, he was informed that he was participating in a learning experiment in which a number of tests of memory would be used. Following this introduction he learned the list of 15 syllables to a criterion of one perfect trial. This was accomplished by exposing the set of cards with the paired syllables at the rate of 2.1 sec. per card and by then exposing the stimulus set at the same rate and having the subject write the correct response. This was repeated, alternately showing the paired syllable set and then the stimulus syllable set until the criterion was reached. Subjects were required to write the appropriate response on a clean sheet of paper each time. A 2.1-sec. exposure is a relatively short time for a written response so the association had to be made quickly.

Following learning to the criterion, the subject was dismissed with instructions to return in three days (72 hours). No indication was given that the next part of the experiment would in any way be associated with the paired syllables. The first few subjects were arbitrarily assigned to a control or experimental group, but later, in order to equate the groups, the subjects were divided on the basis of rate of learning following the initial experimental session.

When the subject returned 72 hours later he relearned the list of 15 paired syllables by the same procedure to a criterion of one errorless trial. Up to this point all subjects had been treated alike and the two groups had been equated for learning ability. At this point, however, the procedure for the two groups varied drastically. The control subjects were shown the five black cubes, four of which were spaced about one in. apart in a straight line, the fifth being held in the experimenter's hand. (The experimenter sat opposite the subject.) The subject was told, "Here are five blocks, four on the table and one in my hand. With this block [indicating the one held in the hand] I will tap the others in some order, for example [taps blocks in 1, 3, 2, 4, order], and you are then to do the same thing. Do you understand?" The subjects grasped the idea rapidly and exhibited a great deal of interest in the 'test.' The level of the task was adjusted so that the subject very seldom failed. Although he was not told his results for the reproduction of any individual pattern, at the end of 15 min. he was told, "You did all right." The examiner appeared to the subject to be keeping a record of the results, and the subject could observe this at any time. This was carefully placed with the subject's learning sheets. After the 15-min. tapping test the examiner announced, "We will now continue with another phase of the experiment." The cards were again presented and the number of trials to relearn to the criterion recorded. The entire situation was one of pleasantly toned relations and the subjects usually left in a cheerful mood. Each was instructed to return 72 hours later.

The experimental group was treated quite differently. It had been found by preliminary experiments that the most acute emotional disturbance would be caused if the subject could be prevented from becoming angry at the experimenter. This is in keeping with the findings of Rosenzweig (3) and Rosenzweig and Sarason (4, 5) that individuals who react impunitively in a frustrating situation repress, while those subjects who react extrapunitively or intropunitively tend to use other defense mechanisms. Consequently, the experimenter appeared in the role of a very concerned although stern individual. The directions given were the same as in the control situation. "Here are five blocks, four on the table and one in my hand. With this block [indicating the one held in the hand] I will tap the others in some order, for example [tap 1, 3, 2, 1, 4, 2, 4] and you are to do the same thing. Do you understand?" If the subject tapped the

blocks from left to right as the examiner did, he had to reverse the pattern since the examiner was on the opposite side of the table. On the other hand, if the subject tapped the blocks in the same direction as the experimenter, his own orientation from left to right was reversed. The first block sequence usually contained seven units which the subject almost surely could not reproduce, or if he could, he was so unsure of himself that he readily believed he had not. Either way he was told by the examiner that he had made a very serious error indicative of poor mental powers. After each failure he was informed that he had failed. The examiner kept score of the successes and failures (all failures) on a sheet which began at seven and ended at fifteen as a score, so that the individual being tested could see himself failing repeatedly at the very bottom of the scale. After a number of failures the examiner elaborately added a new bottom to his scale and informed the individual that he was now working at the eight-year-old level. After a few failures most individuals actually showed a spurt of improvement (which was not indicated to them) compensated for by making the patterns more intricate without increasing their length. Following repeated failure, however, the responses became disorganized. At this point the examiner stopped the experiment temporarily and asked the subject to get hold of himself as he was doing worse than anyone the examiner had ever tested. The subject was informed that no one with such a score could ever hope to get through college and great concern was expressed for the fact that he had such a poor memory. This served to distress the individual still further so that at the end of another few minutes most subjects were sitting on the edge of the chair concentrating all of their attention on the task. The examiner, however, regularly informed the subject of failure after each trial. It was by no means uncommon at this point for the individual's performance to deteriorate to such an extent that he was unable to reproduce even a simple four or five unit sequence. In the last few minutes of the 15-minute period it was, for a few subjects, thought advisable to reduce the tension. In this particular experiment, however, no subject had to be informed of the true purpose at this point.

Following the block tapping test the subject was told that his performance was completely unsatisfactory, that his memory was poor, and that his chances of success were nil. *At no point, however, was any mention made of the nonsense syllables previously learned.*

The subject was then told that it was necessary to continue to another part of the experiment and the syllables were then presented for relearning. As was previously mentioned, a 2.1-sec. interval is relatively short for a written response, so that any decrease in the rapidity of association results in a greatly increased number of trials necessary to reach the criterion. This failure in turn further disorganized the response in all of the subjects. Even when the individual reached the criterion he showed none of the accompaniments of a successfully completed task, but seemed to remain under some tension. At the completion of the relearning he was told to return in three days to continue the experiment. It was at this point and in this group that the greatest incidence of subjects lost through failure to return occurred. The data on these subjects were omitted entirely in the final analysis. The direction in which they would have influenced the results had they completed the experiment cannot be determined. It could be predicted that their inclusion might have increased the significance of the results at sessions 3 and 4 since they repressed so readily; or if the repression were difficult to remove the significance of the statistics at sessions 5 and 6 might have been reduced.

Seventy-two hours later individuals from both groups returned. The control group was treated exactly as it was the second day, that is, there was first a relearning of the syllables followed by a 15-min. session of block tapping, either neutral or mildly pleasant in tone, again followed by a relearning of the syllables.

The experimental group was treated differently. They were first presented with the syllables to be relearned. This seemed to bring back a great deal of the emotion associated with the previous session, although no effort was made to either mention or avoid mentioning it, the pre-test conversation being held to a minimum. After the relearning, the blocks were again produced but at this point the experimenter restructured the situation. The subject was introduced to the test by very easy stages and was allowed to succeed at every trial. Hope dawned! The experimenter then gave the subject a set of fake norms he had prepared showing an entirely different picture, and told the subject that the experiment was one in motivation, and that only very superior subjects were chosen to participate. It was amazing to see the change in attitude on the part of the subject. As far as the experimenter knows not one subject doubted his (the experimenter's) statement. The entire tone of the relationship was altered. The subject was cooperative, happy, and in some cases almost euphoric. Although it was not possible to keep a genuine record of the individuals' block tapping scores, there is little doubt that the individuals in this group exceeded the control group at this

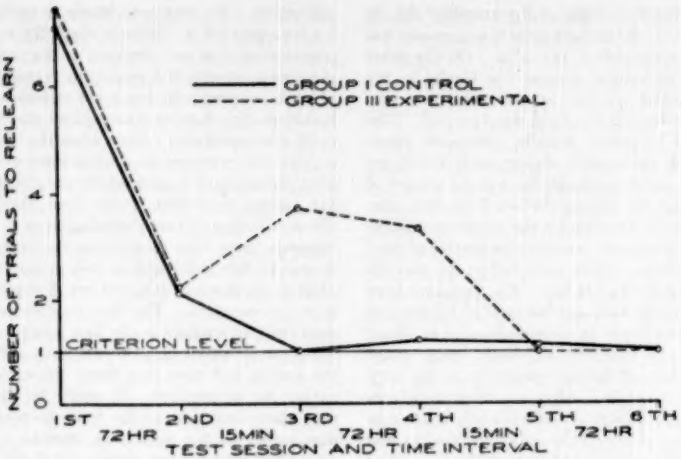

FIG. 1. Experiment I. Mean number of trials to relearn 15 semi-nonsense syllables at each test session

point. Following 15 min. of block tapping (all successes) the next syllable relearning test was administered. The subjects were eager, uninhibited and, as the results show, much better in performance. Seventy-two hours later the individuals from both groups returned for a last test session which lasted only a very few minutes.

Results and discussion.—The results are presented in Figs. 1 and 2 and in Tables I and II. Fig. 1 [3] presents the data in terms of the number

[3] For detailed tables of the data of these experiments, order Document 2710 from American Documentation Institute, 1719 N St., N.W., Washington 6, D. C., remitting $0.50 for microfilm (images one in. high on standard 35-mm. motion picture film) or $1.20 for photocopies (6 × 8 in.) readable without optical aid.

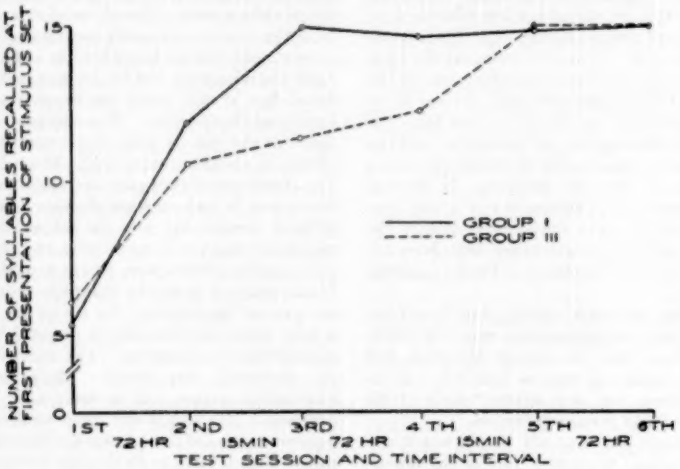

FIG. 2. Experiment I. Mean number of syllables recalled at first presentation of the stimulus set at each test session

TABLE I

EXPERIMENT I. THE *t*- AND *P*-VALUES FOR DIFFERENCES BETWEEN THE MEANS OF EXPERIMENTAL AND CONTROL GROUPS RELEARNING 15 SEMI-NONSENSE SYLLABLES

	Test Session					
	1	2	3	4	5	6
t	0.26	0.33	6.43	4.47	1.00	0.00
P	.40	.37	.0001	.001	.17	—

TABLE II

EXPERIMENT I. THE *t*- AND *P*-VALUES FOR THE DIFFERENCES BETWEEN THE MEANS OF THE CONTROL AND EXPERIMENTAL GROUPS FOR THE NUMBER OF SYLLABLES RECALLED ON THE FIRST TRIAL OF EACH SESSION (RELEARNING OF 15 SEMI-NONSENSE SYLLABLES)

	Test Session					
	1	2	3	4	5	6
t	0.61	1.20	5.00	4.17	1.00	0.00
P	.28	.13	.0004	.001	.17	—

of trials to learn to the criterion, while Fig. 2 presents the same data in terms of the number of syllables recalled at the first presentation of the stimulus set at each learning session.[4] Tables I and II are presented in terms of *P*-values, the probabilities that the measured differences are due to chance factors. As can be seen from the figures, the two groups are approximately equal for Session 1 and Session 2. Session 3, which occurred immediately after the first block tapping session, is the first point at which there is a reliable difference. At this point $t = 6.43$. The probability of a difference this large, in the same direction, resulting from chance is less than one in 10,000. These findings are in keeping with those obtained by Sears (6) and McClelland and Apicella (2). Sears found that failure at card sorting affected the learning of nonsense syllables, and McClelland and Apicella in a very ingeniously designed experiment found that verbal abuse could inhibit the demonstration of learning but that this learning became apparent when the inhibition caused by the abuse was removed. The McClelland and Apicella experiment used card sorting as the original learning and success at pursuitmeter performance as a means of removing inhibition. They explain the improvement in performance in terms of a decrease in anxiety over anticipated failure. In their experiment there is, however, no evidence to warrant a conclusive explanation in terms of repression. That is, the difference between the groups could just as readily be attributed to differences in attitude, set, motivation, or attention, and not to an active repression.

At the next relearning test (4th session) which was administered at the beginning of the third experimental day, there is an absolute difference of 2.1 trials which has a probability of less than eight in 10,000 of being due to chance. Although there is no evidence that the result is not due to forgetting or differences of motivation, such an explanation seems improbable. The results of the fifth session indicate that the removal of the ego threat by success resulted in greatly improved recall. The results of the sixth session indicate that the groups have remained equivalent in performance. The number of subjects at this point was, however, reduced as this last session was incor-

[4] Although only two groups were used in the experiment, the experimental group will be referred to as Group III so that the results of this experiment may be directly compared with results in an experiment to follow.

porated into the design after the experiment had begun.

Taken as a whole it would appear that, in keeping with the original purpose of the experiment, a technique has been found which can be used to induce, in the laboratory under controlled conditions, an emotional experience which fulfills all of the requirements of a Freudian repression sequence. As was pointed out, however, the differential at each stage might be explained in some other manner. For instance, the difference at the third session might be attributed to motivational factors. Although no measure of motivation was attempted, it seemed to the experimenter that the experimental subjects were much more highly motivated than were the members of the control group at this point. A similar point could be made for the 4th session, although here again subjective evaluation would tend to indicate otherwise. Also at the 4th session the experimental group had had on the average more practice than the control group and the passage of time could have served to dissipate any disorganization of response caused by anger or other emotional disturbance. At this point, unless there were repression, it would seem to make sense to predict that the experimental group would either equal or surpass the control group.

Unfortunately, the control group had reached a limit of its measurable learning at the third session, so that it was impossible to determine whether overlearning was taking place. If such were the case, a more delicate measuring instrument might show a difference between the two groups. The results of the fifth session are in keeping with those of McClelland and Apicella (2), who would explain the results in terms of a delayed demonstration of previous learning. They would attribute this to reduced anxiety over anticipated failure. The last relearning shows results which would have been predicted in keeping with any of the above explanations, all subjects having reached the final learning period in theoretically the same emotional condition.

It should be noted that there are several factors which would tend to distort the real picture. One mentioned previously is that the control group reached a criterion level too soon, with the result that no spread was possible. Another is, that the subjects who were in the experimental test group and who 'forgot' to appear for the test on the third day, quite possibly depressed the results. The very fact that they did not appear might indicate that repression had advanced to such a stage that they literally did not remember the appointment! An alternative explanation, of course, is that they remembered only too well and quite consciously avoided returning to have their egos further deflated.

Individual Experiment II

Relearning 30 Semi-Nonsense Syllables

Although the results from the first experiment in this series were statistically clear cut, alternative interpretations of the data were not excluded. This made it desirable to repeat the experiment with further controls. While the fundamental design of the experiment was not altered, the following changes were made: 1. introduction of a group in which failure on the tapping test was designated as specific to tapping; 2. introduction of a group in which failure was designated by the experimenter to apply to both the tapping and syllables test; 3. lengthening of the list of syllables

from 15 to 30 pairs; 4. reduction of the interval between successive testings from three to two days.

The following four groups of 10 subjects each were used in this experiment.

Group I. The control group which suffered no experimentally induced failure. This group is equivalent to the control group in Experiment I.

Group II. A group in which failure at the tapping test was designated by the experimenter as specific to that task. These individuals were told that they had done well on the syllables test. The reason for introducing this group was to determine if failure *per se* could account for the results obtained in the previous experiment. If so, the results for this group should resemble the results of the experimental group in Experiment I. On the other hand, if the induced failure remained specific to the situation, it should not affect the associated material, namely, the syllables test, so the relearnings for the group should then approximate those of Group I. A further possibility was that the emotions associated with the failure designated as specific to the block test might show a partial generalization and cause the results of this group to fall midway between Groups I and III. Another possibility was that such specific failure might facilitate relearning (as compensation) so that Group II would surpass Group I.

Group III. This is the same as the experimental group (Group III) of Experiment I. Failure was induced in the block tapping test and by implication into any other activity the individual might undertake, although no mention of the syllable test was made in connection with the block test.

Group IV. This was a group in which the subjects were told they failed during both the tapping test and the syllables test. It was felt that such a procedure would show whether or not failure irradiating from a single type of material would be equivalent to complete failure on all types. If the performance of this group were equivalent to Group III, then it would seem reasonable to interpret the poor performance of Group III to their generalization of failure to the syllables test. If, on the other hand, being informed of failure in both situations was even more traumatic to the subject, relearning could be expected to show more deterioration. In any case, it was felt that the results would cast more light on the mechanism of repression.

Material.—The material for this experiment was the same as for Experiment I with the exception of the syllable list which was increased to thirty pairs.

Subjects.—The subjects for this experiment were 40 college students of both sexes. All were volunteers, although 20 were paid for their time and 20 were not. Altogether, 48 individuals were used, but for various reasons eight did not complete the experiment. The results of these subjects are not reported. There was no duplication of subjects from the previous experiment in any subsequent experiment.

Procedure.—All subjects learned the list of 30 syllables to a criterion of one perfect trial. This was accomplished as in the previous experiment. The groups were equated through the beginning of the second session by the same method as used in Experiment I.

The procedure was the same for all groups through the relearning experience at the beginning of the second test day. Beyond this point the procedure varied in the different groups.

In the block tapping test the experimenter sat opposite the subject. Four blocks were arranged one in. apart in a row. The experimenter said, "Here are five blocks, four on the table and one in my hand. With this block [indicating the one held in the hand] I will tap the others in some order. For example [taps 1, 3, 2, 4 or 1, 3, 2, 1, 4, 2, 4 depending on the group], and you are to do the same thing. Do you understand?"

Group I was given 15 min. of block-tapping exercises. Members of this group were not told their results, but the patterns were kept at such a low level of difficulty that it was obvious to the subject that he was succeeding very well. In addition, he could see, if he wished to, that the examiner was recording success after each reproduced pattern. At the end of the 15-min. period he was told, "You did all right," and a relearning test was administered. He was then told to return in 48 hours. On the third test day the syllables were again learned to the criterion (4th session) and the procedure of the second day duplicated. There was a 15-min. block-tapping period followed by relearning (5th session). Forty-eight hours later each subject returned for the 6th and last session.

The individuals in Group II were treated differently. The first trial at block tapping was on a difficult seven-unit pattern (1, 3, 2, 1, 4, 2, 4) which was almost certain to be beyond the subject's capacity. In any case, even if the subject succeeded, he readily accepted the suggestion that he had failed. As the examiner was across the table from the subject, if the subject tapped the blocks in the same order, his left-right orientation was opposite to that of the examiner and if he changed the orientation so that he also tapped from left to right, the order of the blocks was reversed. Either way he was informed that he had committed a very stupid error. "I am

very disappointed in you. Your achievement on the nonsense syllables was so good I am very amazed at your complete lack of ability on this test." After each trial (all were failures) the examiner informed the subject he had failed. A careful score was kept showing repeated failures below the bottom of a scale extending from 7 to 15. After a number of failures, the experimenter informed the subject that he was performing at the level of an eight-year-old, but that it was surprising considering his accomplishment on the syllables test. At this point the examiner stopped the experiment, added a new bottom to the scale and told the subjects to get hold of themselves. The results were usually an improved block performance which seemed to deteriorate very little even though they were repeatedly told that they had failed. Although great concern was shown by the experimenter over the subjects' failure at this task, the subjects, although in most cases worried, seemingly did not get really upset about the situation. Several suggested that they had always had poor mechanical ability and all readily admitted that telephone numbers were easily forgotten. It seemed that the feeling of having succeeded on the syllables acted as a strong ego supporting factor so that apparently little emotional trauma was experienced. At the end of 15 min. a relearning test was administered (3rd session). The subjects seemed very glad to get back to this type of material, as was indicated by such remarks as "This is my meat; I'll really knock this off." These subjects seemed to have reached the stage of accelerated performance noted in the failure group of the previous experiment (see Experiment I), but it did not deteriorate as the performance of individuals in that group did after a few trials with repeated failure. Forty-eight hours later a third relearning test (4th session) was given followed by a 15-min. block-tapping session. At this point the individuals were allowed to succeed and were informed of their success. Surprisingly enough, several insisted they could not do the block-tapping test and that their mechanical aptitude had always been poor. They even seemed reluctant to give up the dubious distinction of being the only person to ever make such a low score. After this 15-min. session another relearning test (5th session) was given and the subjects were asked to return in 48 hours. At the last session, which lasted only a few minutes, the syllables were again learned (6th session) to a criterion of one perfect trial.

The subjects in Group III were treated exactly the same as the individuals of the failure group in Experiment I. The difficulty of the task was immediately set above their capacity. They were first informed of their stupidity in orienting themselves wrong in tapping the blocks. They failed on every tapping pattern and were informed of their failure. Care was taken to keep an obvious record, and great concern was shown for their lack of memory ability. The subjects responded first with renewed effort and then as failures continued, the behavior became tense and disorganized which resulted in such deteriorated performance that in one case the individual was incapable of tapping three of the four blocks in the proper order. It seemed to the experimenter that the subjects were very highly motivated at this point. The disorganized behavior was fostered by the examiner who, while seeming to be sympathetic, did everything he could to instill a feeling of absolute and complete inadequacy in the subject. Nothing was omitted in the attempt to make the situation as traumatic as possible. As the aim was to make the situation emotionally unbearable for the individual, the specific type of disparagement resorted to depended to a great extent on the reaction of the subject. All were informed of their stupidity, inadequacy, emotional instability, and lack of both concrete and abstract intelligence. No reference was made at any time, however, to the syllable test during the block tapping test. After this a relearning test (3rd session) was administered and no further reference to failure was made after the block test was finished. The subjects had for the most part become so disorganized that they failed on the first trial. This failure seemed to increase the state of tension to such a point that repeated failures were the rule. Following this test the subject was instructed to return in 48 hours to continue the experiment.

On the third experimental day, each subject was again given a relearning test (4th session). At this time the very sight of the examiner seemed to cause a certain amount of uneasiness, and when the subjects were informed that the first task would be to relearn the syllables, this uneasiness increased. Following this the blocks were produced again and the subject given very easy patterns to reproduce. As more and more of these were successfully reproduced, the tension dissipated and the subject became more relaxed and accuracy increased. At this point it was explained that the experiment had been partially one of motivation and that they had been chosen because of their superior ability. A bogus set of norms was exhibited and the subject was allowed to succeed no matter what he did. Although no legitimate records could be kept of block tapping ability owing to the necessity of falsifying the record during experimental sessions, it seems certain that objective measurement would show that the group at this period reached a maximum of achievement. Following the 15-min. period, a relearning test

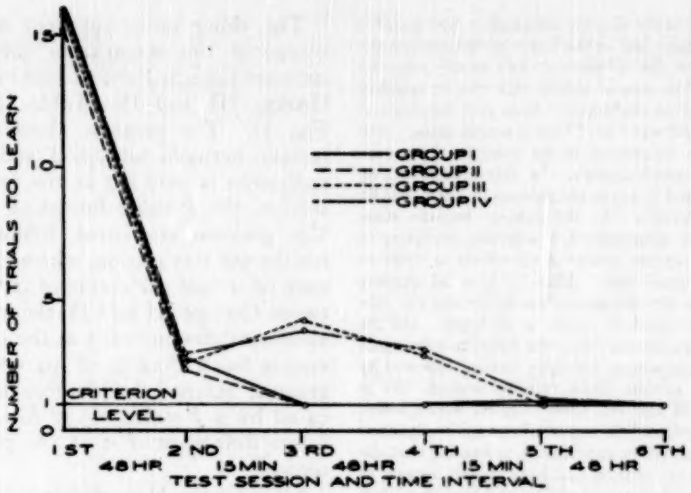

FIG. 3. Experiment II. Mean number of trials to relearn 30 semi-nonsense syllables at each test session

(5th session) was again administered. Forty-eight hours later the subjects returned for the 6th and final session.

The individuals in Group IV were treated the same as the subjects in all other groups up to the time that the first tapping pattern test was given. After the presentation of the first tapping pattern they were informed not only that their first tapping pattern was incorrect, but that this could have been predicted because of their poor showing on the syllables test. Following this, the procedure observed in Groups II and III was

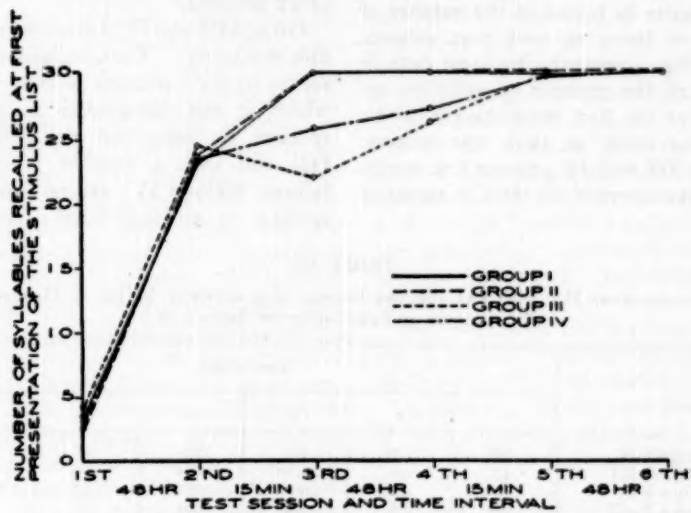

FIG. 4. Experiment II. Mean number of syllables recalled at the first presentation of the stimulus set at each test session

followed fairly closely, although it was possible to add more fuel to the flame by mentioning the failure on the syllables as well as the subject's general low mental ability and specific inability to do well on the blocks. After each tapping test the subject was told, "That is wrong again," with as much resignation in his voice as the experimenter could register. In this group nothing was spared to make the subject feel emotionally uncomfortable. As the subject became more and more disorganized, it was only necessary to let him see the results of his efforts in order to further upset him. After 15 min. of tapping patterns, the 3rd session was given and the subject was asked to return in 48 hours. On the third experimental day, the subjects were again given a relearning test (4th session) followed by another 15-min. block tapping session. As in Groups II and III, these subjects were allowed to succeed, and the experimenter again explained the nature of the experiment, indicating that the subjects in question had been chosen because of their superior mental ability and that they were really doing very well. If possible, the members of this group showed even more relief than those in Group III. After this a relearning test (5th session) was given and the subject was asked to return in 48 hours. At this time the 6th and last relearning was done by the subject.

Results and discussion.—The results are indicated in Figs. 3 and 4, and in Tables III and IV. Figure 3 presents the results in terms of the number of trials to learn at each test session, while Fig. 4 presents the same data in terms of the number of syllables recalled at the first presentation of the stimulus cards at each test session. Tables III and IV present the statistical treatment of the data in terms of *P*.

The thing most apparent at first glance is the remarkable similarity between Groups I and II and between Groups III and IV (Table III and Fig. 3). The greatest absolute difference between Groups I and II at any point is only 0.7 at the first test session, the *P*-value for which is .33. The greatest statistical difference is for the 5th test session, where a difference of .1 had a *P*-value of .17. Between Groups III and IV the greatest absolute difference of .3 at the 4th test session has a *P*-value of .39, while the greatest statistical difference is indicated by a *P*-value of .27 for an absolute difference of .1 at the 5th test session.

Since Group II so closely resembles Group I, it would seem reasonable to assume that failure specific to a single, specifically defined situation offers very little threat to an individual's security. An alternate explanation is that being informed of success on syllable reproduction counteracted the effect of being informed of failure at block tapping.

Groups III and IV also show remarkable similarity. The conclusion which seems to be indicated is that failure which is not designated as specific spreads to associated tasks (Group III) and that a number of specific failures (Group IV) are no more disruptive to an individual's perform-

TABLE III

EXPERIMENT II. *P*-VALUES FOR THE DIFFERENCES BETWEEN MEANS OF GROUPS RELEARNING 30 SEMI-NONSENSE SYLLABLES

	Test Session					
	1	2	3	4	5	6
Groups I–II	.33	.29	.50	.50	.17	.50
Groups I–III	.47	.41	.0001	.001	.50	.50
Groups I–IV	.50	.42	.003	.001	.27	.50
Groups II–III	.36	.36	.0001	.001	.17	.50
Groups II–IV	.35	.26	.003	.001	.08	.50
Groups III–IV	.48	.35	.39	.38	.27	.50

TABLE IV

EXPERIMENT II. *P*-VALUES FOR THE DIFFERENCES BETWEEN THE MEANS OF GROUPS FOR THE NUMBER OF SYLLABLES RECALLED ON THE FIRST TRIAL OF EACH SESSION (RELEARNING OF 30 SEMI-NONSENSE SYLLABLES)

	Test Session					
	1	2	3	4	5	6
Groups I–II	.17	.39	.50	.50	.17	.50
Groups I–III	.27	.25	.0001	.0002	.50	.50
Groups I–IV	.36	.24	.0002	.0001	.43	.50
Groups II–III	.07	.35	.0001	.0003	.17	.50
Groups II–IV	.23	.42	.0002	.0001	.39	.50
Groups III–IV	.16	.28	.09	.07	.43	.50

ance than a failure generalized from a specific situation. It could reasonably have been expected that the individuals in Group IV would have done measurably less well than the individuals in Group III. This, however, did not occur.

The experiment on the whole is subject to many of the same criticisms leveled at Experiment I, namely, that the control group (Group I) reached a level of proficiency by the 3rd test session which made it impossible to measure their learning past this point and consequently impossible adequately to compare the groups. Another possibility is that the individuals used as subjects were not representative of college students in general. In spite of the fact that one half of the subjects were paid for their time, the greatest single factor determining their volunteering was the hope of later discussing personal problems. Next, although the overall picture supports the hypothesis that an actively induced repression can be removed with a consequent revival of memory traces, an analysis at any one point leads to several possible explanations. For example, at the 3rd test session the results, although certainly significant (Diff. I–III = 3.2, $P < .0001$; Diff. I–IV = 2.9, $P < .003$; Diff. II–III = 3.2, $P < .0001$; Diff. II–IV = 2.9, $P < .003$), can be explained in terms other than repression. The subjects in Groups III and IV may have been emotionally disorganized, or may have just quit trying. This last explanation from the examiner's viewpoint does not seem very likely, however, as the individuals in Groups III and IV appeared to be much more highly motivated than the individuals in the other groups. The differences at the 4th session, which are comparable to those of the 3rd session, are open to the same interpretation (Diff. I–III = 1.9, $P < .001$; Diff. I–IV = 2.1, $P < .001$; Diff. II–III = 1.9, $P < .001$; Diff. II–IV = 2.1, $P < .001$).

The results for the 5th test session are in keeping with those obtained by McClelland and Apicella (2). At this point the added practice should have resulted in superior performance for Groups III and IV as compared with Groups I and II. Unfortunately the measuring instrument was not delicate enough to discriminate this possible difference.

The results of the 6th test session are in keeping with any of the previously mentioned explanations.

Apparently the decrease from 72 to 48 hours between experimental days had little effect (compare Figs. 1 and 3).

In any attempt to test experimentally any of the psychoanalytic mechanisms, the experimenter is faced first

with the problem of defining the concept and second of designing an experiment to test adequately the concept. It is always possible that somewhere in this process the true meaning of the original concept has been lost. It seems, however, that the concept of repression proper or after-expulsion can be defined as an active process which forces once conscious material from consciousness without destroying the memory for the material, which memory may be restored following the removal of the inhibiting factor. The present design seems adequate to test this hypothesis. Although alternative explanations are possible, it seems reasonable to conclude that the results indicate that an experimental analogue of repression has been produced in the laboratory.

Summary and Conclusions

Two experiments have been conducted to test the hypothesis that repression proper is an active process which inhibits memory for previously known material which has become unacceptable to the individual, but that the memory for the material may be restored by reducing the ego threat associated with the original material. The neutral material consisted of 15 and 30 semi-nonsense syllables respectively for Experiments I and II. Induced failure in reproducing a series of tapping patterns using five one-in. black Knox cubes constituted the ego threat, while success at the same task served to remove the threat. Neutral experience at tapping was the interpolated material in the control group. The results indicate:

1. That induced failure at a task, when not indicated as specific to that task, serves to reduce the ability to recall previously known material which has become associated with the failure task.

2. That this reduced ability lasts for a period of time (three days in Experiment I and two days in Experiment II).

3. That induced success at the same associated task serves to increase the ability to recall the original material.

4. That implied failure is as disruptive to memory as specific knowledge of failure.

5. That failure indicated as specific to a given task has no measurable effect on the ability to perform other tasks.

6. That although other explanations in terms of motivation, interest, or attitude cannot be ruled out, the experiment fulfills the requirements of an analogue of Freudian repression.

(Manuscript received August 11, 1949)

References

1. Flanagan, D. E. The influence of emotional inhibition on learning and recall. Unpublished Master's thesis, Univ. of Chicago, 1930.
2. McClelland, D. C., & Apicella, F. S. Reminiscence following experimentally induced failure. *J. exp. Psychol.*, 1947, 37, 159–169.
3. Rosenzweig, S. An outline of frustration theory. In: J. McV. Hunt (Ed.), *Personality and the behavior disorders*. New York: Ronald Press, 1944.
4. Rosenzweig, S., & Sarason, S. An experimental study of the triadic hypothesis: reaction to frustration, ego-defense, and hypnotizability. I. Correlational approach. *Character and Pers.*, 1942, 11, 1–19.
5. Sarason, S., & Rosenzweig, S. An experimental study of the triadic hypothesis: reaction to frustration, ego-defense, and hypnotizability. II. Thematic apperception approach. *Character and Pers.*, 1942, 11, 150–165.
6. Sears, R. R. Initiation of the repression sequence by experienced failure. *J. exp. Psychol.*, 1937, 20, 570–580.
7. Zeller, A. F. An experimental analogue of repression. I. Historical summary. *Psychol. Bull.*, 1950, 47, 39–51.

THE EFFECT OF DIFFERENT TYPES OF PRELIMINARY ACTIVITIES ON SUBSEQUENT LEARNING OF PAIRED-ASSOCIATE MATERIAL [1]

BY LELAND E. THUNE [2]

From the Psychological Laboratory, State University of Iowa

Introduction

The present study is concerned with the effects of different amounts and types of certain preliminary activities upon the immediately subsequent learning of paired-associate material when: (1) the total amount of prior practice at the time of the learning of the test list is held constant, (2) the time interval between the cessation of the preliminary activity and the initiation of learning on the test list is brief (only a few minutes), and (3) the preliminary activities are of such a nature as to produce a minimum of proactive inhibition in the learning of the test list. There is little material in the literature which bears closely upon this problem. The most relevant studies have been done in the context of warm-up phenomena in motor learning, and the majority of these investigations have been concerned with the efficiency of an already learned or practiced performance after an interpolated rest interval rather than with the learning of new materials.

Warm-up phenomena in studies involving relearning or continuation of pre-rest practice

In a recent theoretical analysis of the motor learning situation, Ammons (1) assumes that warm-up decrement (the decrement in efficiency attributable to the need to warm-up) originates from loss of set during rest, and that set consists of postural and other muscular responses which are acquired during pre-rest practice and which facilitate subsequent performance. Factors such as amount of pre-rest practice and length of interpolated rest interval, which act to determine the number and strength of the particular postural and muscular responses carried over to the post-rest trials, thereby determine the magnitude of the warm-up decrement.

In two recent discussions of warm-up phenomena, Irion (2, 3) accepts Ammons' identification of the warm-up factor with the reinstatement of set lost during the rest interval, and then goes on to theorize regarding the conditions under which loss of set might occur. Loss of set is assumed to be some function of the nature of the activity which *S* engages in during the rest or retention interval. According to Irion's hypothesis, activities which disrupt the set acquired during original learning should be followed by a greater warm-up effect than would follow interpolated activities which tend to favor the maintenance of the original set. Irion's (2) conceptualization of the warm-up factor leads him to differentiate loss of retention into two classes of factors namely, (1) loss of retention attributable to altered set, and (2) loss of retention attributable to a residual factor which would include factors such as altered stimulating conditions and competition from incompatible responses.

Irion (2) speculates briefly upon the manner in which change in set may operate to produce a loss in retention. He identifies certain factors which are thought to determine the efficiency with which the responding organism receives stimulus cues from the environment and

[1] Based on a dissertation submitted to the faculty of the Department of Psychology of the State University of Iowa in partial fulfillment of the requirements for the Ph.D. degree. The author is greatly indebted to Professor Kenneth W. Spence who directed the investigation.

[2] Now at Vanderbilt University.

is able to produce behavioral responses which are appropriate to those cues. An additional factor, that of stimulus generalization, is also set forth. Irion notes that:

> Part of the stimulus complex to which the individual has learned to respond may be thought of as originating internally as proprioceptive and kinesthetic stimulation. To the extent that this component of the stimulus context is altered between the time of original learning and the time of recall, the strength of the response tendency may be reduced by the operation of stimulus generalization and stimulus interaction (2, pp. 22–23).

Irion (2) reported five experiments dealing with reminiscence and retention as a function of the length of the rest interval and of the amount of pre-rest practice. Two of these studies deal with motor learning in a pursuit rotor task and three are concerned with verbal learning in a memory drum situation. The motor studies measured retention as a joint function of the amount of pre-rest practice and of the length of the retention interval. These two experiments demonstrate warm-up effect to be an increasing function of both variables, a result which is in agreement with the theoretical formulations and empirical findings reported previously by Ammons (1).

Two of Irion's verbal learning studies were likewise designed to measure retention as a function of the amount of pre-rest practice and of the length of the retention interval. Again, warm-up effect was found to be an increasing function of both of these variables. Thus, the results obtained by Irion in his verbal learning studies are in agreement with the motor learning data of Ammons and Irion reported above, and to that extent serve to establish the generality of a positive relation between warm-up effect and both amount of pre-rest practice and length of retention interval.

The last of Irion's experiments was conducted for the purpose of investigating the relation between warm-up effect and loss of set. If loss of set to perform the activity under consideration bears a relation to the magnitude of the warm-up effect, then any situation or activity which would tend to reinstate the original set should, if engaged in just prior to the relearning of the original activity, tend to reduce the magnitude of the warm-up effect and hence to increase retention (performance at recall). As a test of this deduction, Irion ran a group of *S*s under a paired-associate verbal learning condition which called for a brief period of participation in a set-reinstating activity just prior to the beginning of the relearning trials at the end of a 24-hour rest interval. As a set-reinstating activity, Irion used a color-naming task in which *S* sat in front of a memory drum and called out the colors of a series of colored paper strips which replaced, for one trial, the usual paired-associate stimulus and response items. *S* had been instructed to the effect that, "The purpose of the color naming is to get you back into the swing of the learning task" (2, p. 104). The effect of the special instructions plus the one color-naming trial was that of raising the initial level of the relearning performance until it almost equalled the performance level attained by a group which received only one min. of rest between the original and relearning periods but which did not engage in any set-reinstating activity. This experiment of Irion's demonstrates in a striking manner that, following the reinstatement of set, there may be a significant increase in apparent retention.

Warm-up phenomena in the original learning of new material

Many studies have reported on the effects which practice at one activity may have upon subsequent performance on a new activity. The majority of these investigations have been concerned either with the transfer of practice effects or with the phenomena of proactive inhibition. No studies have been reported in which there has been a direct attempt to measure the effects upon the learning of new materials of immediately preceding practice on a different activity when level of practice is held constant and the effects of proactive inhibition are mini-

mized. However, a study reported by Heron (5) provides indirect evidence bearing on this topic.

Heron's investigation involved three experimental conditions each of which entailed the learning of two lists of paired-associate nonsense syllables, with practice on the two lists being separated by a two-min. interval. The conditions of Heron's experiment originally were designed to investigate negative transfer. Under two of the three conditions the two lists learned on any one day were related by using some of the same stimulus syllables in both lists. Heron had thought that by relating certain of the pairs in this A-B, A-K manner, he might obtain a negative transfer or interference effect from the learning of the first list to the learning of the second. No such interference effect was obtained. However, when the experimental data were analyzed without regard to conditions, it was found that the first list learned on the second experimental day required significantly more trials to learn than did the second list learned on the first day. Likewise, the first list learned on the third day required significantly more trials to master than did the second list learned on either of the two preceding experimental days. The control condition, in which the two lists were not related, showed as much facilitation in the learning of the second list as did the other two conditions, thus causing Heron to reject direct positive transfer of specific learned units from the first list to the second list as an explanation for the more rapid learning of the second lists. There was only a slight general practice effect from one day's performance to the next, the first list learned on the third day requiring only slightly fewer trials than the first list learned on the first experimental day. Heron's experimental results led him to conclude that the same warm-up effect was operating in his verbal learning situation, which involved new materials, as had previously been demonstrated in the relearning of motor skills.

Purpose of the present investigation

The purpose of the present experimental investigation was to obtain a more adequate picture of the role of the warm-up factor as it operates during the original learning of verbal materials. Experiment A was designed specifically to measure the effect upon the learning of a new paired-associate list of varying amounts of immediately preceding practice on a different paired-associate list when level of practice is held constant and the effects of proactive inhibition are minimized. Experiment B grew out of Irion's (2, 3) work on the effect of set reinstatement upon retention. An attempt has been made to investigate the effects upon the learning of a new paired-associate list of immediately preceding practice on an activity which does not involve associative learning but which might be expected to instill a response set similar to that which is established during paired-associate learning practice, again with level of practice held constant and the effects of proactive inhibition minimized.

Experimental Procedure

Experiment A

Experiment A was designed to measure the degree to which the learning of a paired-associate list influences the rate with which a second paired-associate list is learned after a 10-minute cartoon sorting interval.[3] A schema of the conditions which prevailed in Experiment A is shown in Table I. Each *S* participated in two experimental sessions and learned three paired-associate lists which have been designated as (1) a Practice List, (2) List A, and (3) List B. The various experimental conditions have been

[3] The design of this experiment was such as to coordinate the study with a related investigation undertaken concurrently by Hamilton (4) in this same laboratory. Condition 10 of the present study is identical with one of Hamilton's conditions. Hamilton served as *E* for 20 of the 25 *S*s being reported under Condition 10 in the present study.

TABLE I

SEQUENCE OF EXPERIMENTAL EVENTS USED TO MEASURE THE EFFECT OF 0, 2, 4, 6, 8, AND 10 PAIRED-ASSOCIATE LEARNING TRIALS UPON SUBSEQUENT LEARNING OF NEW PAIRED-ASSOCIATE MATERIAL

Condition	Day 1			Day 2		
	Practice List	Rest	List A	List A	Cartoon Sorting	List B
0	6 trials	1′	10 trials	0 trials	10′	10 trials
2	6 trials	1′	8 trials	2 trials	10′	10 trials
4	6 trials	1′	6 trials	4 trials	10′	10 trials
6	6 trials	1′	4 trials	6 trials	10′	10 trials
8	6 trials	1′	2 trials	8 trials	10′	10 trials
10	6 trials	—	—	10 trials	10′	10 trials
0′	6 trials	1′	10 trials	—	—	10 trials

identified in a manner such that the number of any condition corresponds to the number of trials on List A which that group received on the second experimental day. Thus, *Ss* assigned to Condition 6 were given six trials on List A on the second day, whereas *Ss* under Condition 0 received no practice on List A on the second day.

The paired lists of words were typed in capital letters on a glazed white cloth tape and were presented on a modified Hull-type memory drum. The stimulus word of each pair was exposed for a total of four sec., the stimulus and response words being jointly exposed during the last two sec. of this interval. Three different serial orders of presentation were employed. Between trials there was an eight-sec. interval during which the drum continued to turn but only blank spaces appeared in the exposure window. *E* sat behind the exposure apparatus, a screen shielding him from the view of *S*.

The experimental design shown in Table I permits a determination of the effect, on the rate of learning List B, of varying amounts (trials) of learning activity on List A on that same day. It should be noted, however, that the total amount of practice up to the time of learning List B was held constant for all conditions. This was achieved by having the total number of trials on List A on Day 1 and Day 2 equal 10. Thus, in Condition 2, eight trials were given on Day 1 and two on Day 2, while in Condition 8, two trials were given on Day 1 and eight on Day 2.

Condition 0′ was introduced into the design of Experiment A in an attempt to evaluate the effect of the 10-min. cartoon sorting interval on the second experimental day upon the subsequent learning of List B. Conditions 0 and 0′ involved identical experimental procedures except that, on the second experimental day, *Ss* under Condition 0 engaged in 10 min. of cartoon sorting immediately prior to the learning of List B, whereas *Ss* under Condition 0′ did not participate in this preliminary activity.

All learning was carried to a trial criterion, and all *Ss* under all conditions received the same amount (number of trials) of practice on any one list. The Practice List was presented for six trials, List A for a total of 10 trials, and List B for 10 trials. Each list was composed of 15 pairs of words. All *Ss* learned the same series of 15 pairs as a Practice List. Likewise, the same List A and the same List B were learned under all conditions. The three lists were of the A-B, C-D, E-F type in which different and unrelated words were used in each of the lists. The Practice List was composed of one-syllable nouns, List A of two-syllable adjectives, and List B of one-syllable nouns. Several factors entered into the choice of these particular types of learning materials. First, in an attempt to minimize the effects of proactive inhibition, it was considered desirable to have the two lists that would be studied on any one day as dissimilar as possible. Second, it was felt that the principal list that was to be used to measure the influence of preliminary learning practice (List B), should be of such a level of difficulty that most *Ss* would achieve a relatively high degree of mastery within the number of trials given. It was thought that this characteristic of the test list might make it a more sensitive measure of the effects of preliminary learning activity than would a more difficult list. The lists were constructed in such a manner as to eliminate, or at least minimize, any formal or meaningful similarities or associations between stimulus and response words within the same list and between the various lists. Each *S* served under only one condition, thus eliminating the necessity of counter-balancing the various lists and conditions.

One hundred and seventy-five *Ss* (73 men and 102 women) served in Experiment A, 25 serving in each of the seven basic conditions. The *Ss* were all undergraduate students in psychology and were predominantly Sophomores and Juniors. None had had previous experience in a memory drum experiment. Each *S* served on two consecutive experimental days. With but a few exceptions, *Ss* were assigned in the order of their appearance to the various experimental conditions taken in rotation. In a few cases scheduling demands necessitated that a condition requiring a certain length session be given at a particular time. In all such cases the decision as to which experimental condition would be used was made prior to the assignment of any particular *S* to that session. Any *S* who failed to achieve at least one correct anticipation during the six learning trials on the Practice List was automatically eliminated and replaced by the next *S* scheduled to appear for his first experimental session. Under all seven experimental conditions, a total of five *Ss* were eliminated because of failure to meet this minimum criterion. All *Ss* under all experimental conditions received the same instructions.

Experiment B

Experiment B was designed to determine the effect of participation in a memory drum task not involving associative learning upon the immediately subsequent learning of a new paired-associate list. Table II shows the schedule of events used in the two conditions of this experiment. Since the procedure for Experiment B parallels closely some aspects of Experiment A, it will not be necessary to discuss its design in as full detail as was required for Experiment A.

The color guessing task referred to in Table II is a memory drum activity which does not involve associative learning. *S* sat in front of a memory drum and attempted to guess, presumably through the use of 'extra-sensory-perception,' which one of five colors (red, yellow, blue, green, or black), would appear in the response space when the shutter lifted. The same stimulus, a large capital X, was used for all colors on all trials. In order to eliminate the possibility of practice at serial learning, the color patches were arranged in five lists with 15 color patches in each list. Each of the five possible colors appeared three times in each list, the order of arrangement of the colors within each list being left entirely to chance.

Experiment B involved paired-associate learning with the same Practice List, List A, and List B as were used in Experiment A. The same total number of trials per list and the same serial orders of presentation were used in both experiments. For Experiment B, the apparatus used in Experiment A was modified to the extent of having two identical memory drums placed side by side, one being used to present the word lists and the other the color guessing lists.

Fifty undergraduate students (31 men and 19 women) in psychology served on two consecutive days as *Ss* in Experiment B, 25 serving in each condition. Two *Ss* were discarded and replaced because of failure to achieve any correct responses on the six trials on the Practice List. The general instructions used in Experiment A were also used in Experiment B. In addition, *Ss* under Experiment B were given special instructions relative to their task in the color guessing situation.

Results

Warm-up from List A to List B in Experiment A

Degree of learning on the Practice List.—All *Ss* received six trials on the Practice List. The mean number of correct anticipations per *S* on the total of six trials for Conditions 0, 2, 4, 6, 8, and 10 are given in Table III. An analysis of variance of the total number of correct anticipations per *S* on the six trials on the Practice List

TABLE II

Sequence of Experimental Events Used to Measure the Effect of 0 and 10 Trials of Color Guessing upon Subsequent Learning of Paired-Associate Material

Condition	Day 1					Day 2		
	Practice List	Rest	List A	Rest	Color Guessing	Color Guessing	Rest	List B
0-cg	6 trials	1′	10 trials	2′	10 trials	0 trials	0′	10 trials
10-cg	6 trials	1′	10 trials	0′	0 trials	10 trials	2′	10 trials

TABLE III

MEAN NUMBER OF CORRECT ANTICIPATIONS DURING ALL SIX TRIALS ON THE PRACTICE LIST AND ALL 10 TRIALS ON LIST A

Condition	Practice List		List A	
	Mean	σ_M	Mean	σ_M
0	15.28	1.83	46.16	4.80
2	16.08	2.76	33.60	3.77
4	16.72	1.81	35.52	3.91
6	15.44	2.68	33.16	4.17
8	17.44	2.22	27.24	4.06
10	18.04	1.81	35.24	4.00

was computed in order to test the initial comparability of the six groups. The obtained *F*-ratio of .247 is clearly not significant for the 5 and 144 *d.f.* involved in this test. Since the hypothesis of initial comparability between groups remains tenable by the above test, any significant differences between groups on the learning of the other lists can be reasonably attributed to factors other than differences in initial ability for paired-associate learning.

Degree of learning on List A.— Table III also gives the mean number of correct anticipations per *S* on the total of 10 trials on List A for Conditions 0, 2, 4, 6, 8, and 10. In interpreting these data, cognizance must be taken of the fact that, although all *S*s received a total of 10 trials on List A, the relative distribution of these trials on the two experimental days varied from one condition to another. Further discussion of the possible significance of the varying levels of achievement attained under the different conditions will be postponed until later.

Degree of learning on List B.— List B was learned 10 min. after completing varying numbers of trials on List A on the second experimental day. Fig. 1 presents data relative to the mean number of correct anticipations per *S* on the total of 10 trials on List B for Conditions 0, 2, 4, 6, 8, and 10. Fig. 2 shows how the mean number of correct anticipations varied on trials 2 through 10 in the learning of List B for each of these same groups.

In order to evaluate the significance of the obtained differences between groups, the learning scores on List B for Conditions 0, 2, 4, 6, 8, and 10 were analyzed by means of an analysis of covariance. Total number of correct anticipations per *S* on all six trials on the Practice List was used as the initial measure, and total number of correct anticipations on the 10 trials on List B was used as the final measure. The obtained *F*-ratio of 3.27 is significant at the one percent level for the 5 and 143 *d.f.* involved in this test. Thus, the null hypothesis, with respect to performance on List B, may be rejected at a high level of confidence for Conditions 0, 2, 4, 6, 8, and 10. It is readily apparent that, under the conditions of the present experiment, the mean rate of learning on List B bears a direct relation to the number of trials on List A completed 10 minutes previously.

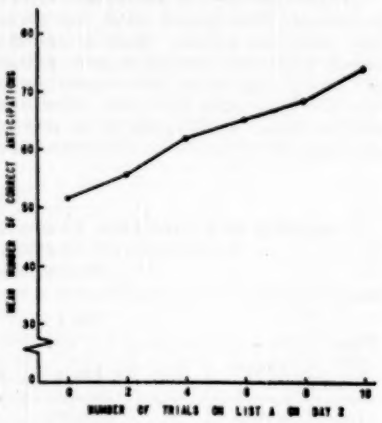

FIG. 1. Mean correct anticipations on the total of 10 trials on List B

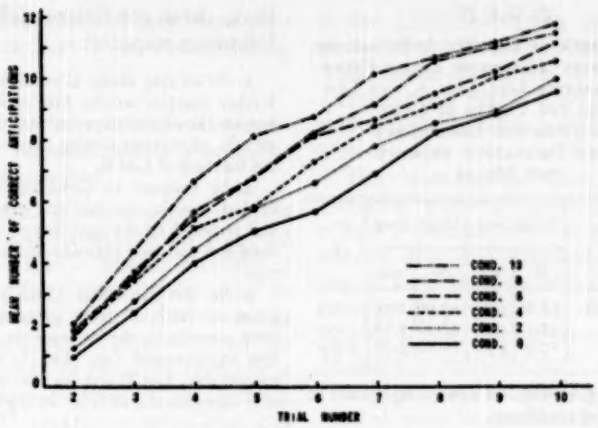

FIG. 2. Mean correct anticipations on trials 2 through 10 on the learning of List B

The effect of cartoon sorting upon the learning of List B

Under Condition 0, Ss received no practice on List A on the second experimental day, but they did engage in 10 min. of cartoon sorting immediately prior to the learning of List B. Condition 0′ was inserted into the design of Experiment A in an attempt to evaluate the effect of this 10-min. cartoon sorting interval upon the immediately subsequent learning of List B. Conditions 0 and 0′ involved the same experimental schedule on the first experimental day, i.e., both groups received six trials on the Practice List followed by 10 trials on List A. On the second experimental day, Ss under Condition 0′ learned List B as their initial task whereas, as has been noted above, Ss under Condition 0 received 10 min. of cartoon sorting as a preliminary activity immediately prior to the learning of List B.

There were no significant differences between Conditions 0 and 0′ in the rates with which any of the three paired-associate lists were learned. A t-value of .108 was obtained for the difference between Conditions 0 and 0′ on the learning of List B. This value lacks significance at even the 90 percent level. Thus it is evident that 10 min. of cartoon sorting in itself had no significant effect upon the rate with which List B was subsequently learned.

Color guessing as a warm-up activity

Experiment B was designed to yield data relative to the effect upon the subsequent learning of List B of 10 trials on the color guessing task when the latter was presented as a set-reinstating activity immediately prior to the learning of List B. The two conditions comprising Experiment B, Conditions 0-cg and 10-cg, both called for six trials on the Practice List followed by 10 trials on List A on the first experimental day, and for 10 trials on List B on the second day. They differed in that Condition 0-cg called for 10 trials of color guessing to be given immediately after the learning of List A on the first experimental day, whereas Condition 10-cg called for 10 trials of color guessing to be given just prior to the learning of List B on the second experimental day.

TABLE IV

MEAN NUMBER OF CORRECT ANTICIPATIONS ON THE TOTAL NUMBER OF TRIALS GIVEN ON THE PRACTICE LIST, LIST A, AND LIST B, AND THE VALUES OF *t* FOR EVALUATING THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN THE MEANS

List	Cond. o-cg		Cond. 10-cg		*t*
	M	σ_M	M	σ_M	
Practice List	17.84	2.76	17.96	2.15	.034
List A	45.12	4.90	47.36	5.18	.292
List B	49.92	5.78	71.36	4.77	2.862*

* For 48 *d.f.* a *t*-value of 2.681 is significant at the 1% level of confidence.

Table IV gives data relative to the mean number of correct anticipations on the total number of trials given on the Practice List, List A, and List B for Conditions o-cg and 10-cg, the difference between the means, and the value of *t* for Fisher's test of the significance of a difference between the means of independent samples. From the data given in Table IV it is apparent that the two groups of *Ss* represented by Conditions o-cg and 10-cg did not differ significantly in initial ability for paired-associate learning as measured by performance on either the Practice List or List A. On the other hand, the *t*-test indicates that the obtained difference between Conditions o-cg and 10-cg on the learning of List B is highly significant.

Possible practice effects from color guessing on Day 1

The schedules of experimental events under Conditions o, o′, and 10-cg were identical through the learning of List A, i.e., on the first experimental day all *Ss* under these conditions received six trials on the Practice List which were immediately followed by 10 trials on List A. After that, these conditions differed in the following respects:

1. *Ss* serving under Condition o received no further practice on the first experimental day, but on the second experimental day were given 10 min. of cartoon sorting immediately prior to the learning of List B.
2. *Ss* assigned to Condition o′ received no further practice on the first experimental day and returned on the next day to learn List B as their first and only experimental activity on that day.
3. *Ss* serving under Condition o-cg were given 10 trials of color guessing immediately after completing the 10 trials on List A on the first experimental day. On the second experimental day, List B was learned as the first and only experimental activity during that session.

This feature of the experimental design makes it possible to determine whether or not participation in the non-associative color guessing activity on the first experimental day resulted in a genuine practice effect which carried over to subsequent performance on List B on the second experimental day. If such a practice effect were present, it should result in a relative superiority of Condition o-cg over Conditions o and o′ in the learning of List B. It has been demonstrated above that the 10-min. cartoon sorting interval on the second experimental day, which distinguished Condition o from Condition o′, had no significant effect upon the subsequent learning of List B.

An analysis of covariance of the total number of correct anticipations on all 10 trials on List B for Conditions o, o′, and o-cg was computed, using total number of correct anticipations on the 10 trials on List A as the initial measure. The obtained *F*-ratio of .93 clearly lacks statistical significance for the 2 and 71 *d.f.* involved in the test. Thus, it is entirely reasonable to retain the hypothesis that participation in the color guessing task on the first experimental day did not result in a general practice effect

which carried over and influenced performance on List B on the following day.

Comparative effectiveness of color guessing and paired-associate practice as warm-up activities

On the second experimental day, Ss serving under Condition 10 received 10 trials on List A and 10 minutes of cartoon sorting immediately prior to the learning of List B, whereas Ss under Condition 10-cg received 10 trials on the color guessing task and a two-min. rest interval before learning List B. If certain assumptions can be shown to be reasonable, it will be possible to evaluate the relative effectiveness of practice at paired-associate verbal learning and practice at non-associative color guessing in influencing the rate with which a new paired-associate list is learned after a brief interval.

The primary assumptions which must be made before the relative effectiveness of these two types of preliminary memory drum activities can be evaluated are: (1) that the groups involved in such a comparison have essentially equal ability for paired-associate learning; (2) that the color guessing activity itself produced no significant general practice effect which carried over to the learning of List B; (3) that the facilitating effects of practice at paired-associate learning carried over a 10-min. cartoon sorting interval with essentially the same effectiveness as the effects of practice at color guessing carried over a two-min. rest interval.

Conditions 10 and 10-cg were quite comparable in initial ability for paired-associate learning as measured by performance on the Practice List. Under Condition 10 a mean of 18.04 correct anticipations was given on the six trials on the Practice List, whereas under Condition 10-cg there was a mean of 17.96 correct anticipations. This mean difference of .08 correct anticipations yields a t-value of .028 which lacks significance at the 90 percent level.

The assumption that the color guessing activity produced no significant general practice effect which carried over and influenced the learning of List B has been shown to be entirely reasonable by the data reported above for the learning of List B under Conditions 0, 0′, and 0-cg.

Data from the present experiment provide no basis for determining whether or not the facilitating effect derived from paired-associate learning practice carried over a 10-min. cartoon sorting interval with essentially the same efficiency as the effects of participation in the color guessing activity carried over a two minute rest interval. However, a fairly satisfactory answer to this question can be obtained from a study by Hamilton (4) undertaken in this same laboratory. Using the same materials and experimental design as were used in Condition 10 of the present study, Hamilton has investigated, among other things, the effects of cartoon sorting intervals of from eight sec. to 40 min. in length upon the immediately subsequent learning of a new paired-associate list. Hamilton's results suggest that the facilitating effects derived from practice on one paired-associate list and acting upon a second paired-associate list are carried over intervals of from eight sec. to 10 min. with but only a minor decrement.

Utilizing the three assumptions noted above, an attempt was made to evaluate the relative effectiveness of practice at non-associative color guessing and at paired-associate verbal learning in influencing the rate with which a new paired-associate list is learned after a brief interval. This was done by means of an analysis of covariance of the total number of correct anticipations on all 10 trials

on List B for Conditions 10 and 10-cg. Total number of correct anticipations per *S* on all six trials on the Practice List was used as the initial measure. The obtained *F*-ratio of .11 gives us no basis for rejecting the hypothesis of equality between Conditions 10 and 10-cg on the learning of List B. Thus, we may tentatively conclude that, under the conditions of the present experiment, paired-associate verbal learning and non-associative color guessing are approximately equal in effectiveness in facilitating subsequent learning of a new paired-associate list after a brief interval.

Data relative to the rate with which List B was learned under Conditions 0′, 10, 0-cg, and 10-cg are shown graphically in Fig. 3. It will be remembered that Conditions 0′ and 0-cg are identical except that Condition 0-cg called for 10 trials of color guessing immediately after learning List A on the first experimental day.

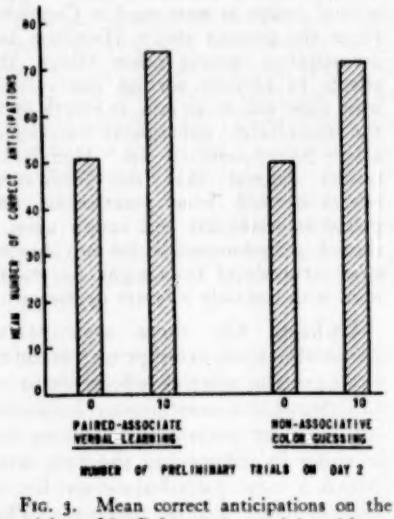

FIG. 3. Mean correct anticipations on the 10 trials on List B for groups receiving either 0 or 10 paired-associate learning or color guessing trials as a warm-up activity prior to learning List B

Both of these conditions differed from Conditions 10 and 10-cg in that under the former conditions no preliminary memory drum practice was given immediately prior to the learning of List B on the second experimental day.

Warm-up effects in the learning of List A

Thus far in the treatment of the results, data have been presented relative to the extent to which paired-associate learning practice on List A or participation in the color guessing activity affected the rate with which List B was learned after a brief interval. However, the designs of Experiments A and B were such that the effect of participation in a preliminary memory drum activity upon subsequent learning of a paired-associate list may also be determined by the effect of the Practice List upon the subsequent learning of List A.

All *S*s serving under Conditions 0, 0′, 0-cg, and 10-cg received 10 trials on List A immediately after receiving six trials on the Practice List on the first experimental day. *S*s under Condition 10 likewise had six trials on the Practice List on the first day, but they had no further memory drum practice until List A was learned for 10 trials as the first activity on the second experimental day. For the purpose of the present analysis, the data of the 25 *S*s who served under each of Conditions 0, 0′, 0-cg, and 10-cg have been combined, the whole set being treated as one group of 100 *S*s. Since all *S*s under these four experimental conditions received exactly the same treatment up through the learning of List A, their pooled results should provide a better estimate of the effect of the Practice List upon the immediately subsequent learning of List A than could be obtained from any one group alone.

A mean of 44.69 correct anticipations was given on the total of 10 trials on List A by the 100 *Ss* serving under Conditions 0, 0′, 0-cg, and 10-cg, all of whom received six trials on the Practice List just prior to the learning of List A. The 25 *Ss* under Condition 10, who received no preliminary practice immediately prior to the learning of List A on the second experimental day, gave a mean of 36.24 correct anticipations on the 10 trials on List A. Thus, *Ss* who received six trials on the Practice List immediately prior to the learning of List A attained 23 percent more correct responses on the learning of List A than did *Ss* who learned List A without having participated in any preliminary memory drum practice on that same day. The significance of this difference was evaluated by means of an analysis of covariance in which total number of correct anticipations on the six trials on the Practice List was used as the initial measure. The obtained *F*-ratio of 4.84 falls between the values required for significance at the five and one percent levels for the 1 and 122 *d.f.* involved in this test. Thus, the hypothesis of no difference between Conditions 0, 0′, 0-cg, and 10-cg all taken together, and Condition 10 in the rate with which List A was learned may be rejected at a reasonably high level of confidence.

Locus during the learning of List B within which warm-up practice has its effect

In Fig. 2 are shown the curves of acquisition for trials 2 through 10 on the learning of List B under Conditions 0, 2, 4, 6, 8, and 10. The slopes of these curves, taken at various points along their extent, may be used as an index of the comparative rates with which List B was being learned at those points under the various experimental conditions. By means of this type of analysis, it should be possible to determine whether the facilitating effects of preliminary memory drum practice act in a constant manner throughout the course of the learning of a new list, or whether the beneficial effects act principally during one particular stage or level in the acquisition of the new list.

For purposes of the present analysis, the learning curves on List B for Conditions 0′, 0, 2, 4, 6, 8, and 10 were divided into thirds, each third including three anticipation trials. The slope within each third was measured in terms of the mean increase in the number of correct anticipations during that one-third of the learning. Thus, the slope for the first third was measured in terms of the mean increase in the number of correct anticipations between trials 1 and 4. In the same manner, the slope for the middle one-third involved the differ-

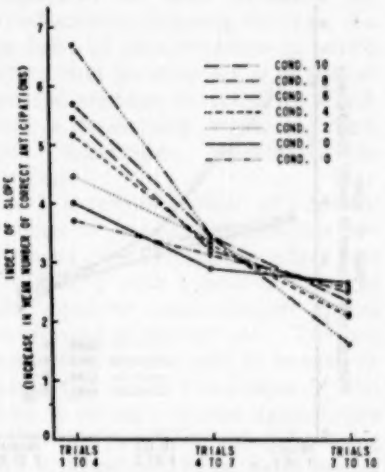

FIG. 4. Analysis of the slopes of learning curves for List B when slope is measured in terms of the increase in mean number of correct anticipations from trials 1 to 4, 4 to 7, and 7 to 10. (Based on empirical data.)

ence between trials 4 and 7, and the slope for the last one-third involved the difference between trials 7 and 10.

Fig. 4 presents curves for Conditions 0′, 0, 2, 4, 6, 8, and 10 when this index of slope (mean increase in the number of correct anticipations) is plotted against the three divisions in the learning of List B which are described above. When the data from Conditions 0-cg and 10-cg were treated in the same manner, curves were obtained which were similar in form to those of Conditions 0 (or 0′) and 10, respectively, in Fig. 4.

This slope analysis based upon the empirical data strongly suggests that the facilitating effects of preliminary memory drum practice act primarily during the first few trials in the learning of the new list. However, these curves have been derived from data based upon learning scores on only three of the nine anticipation trials, namely trials 4, 7, and 10. Any chance irregularities which characterize these three points are reflected strongly in the slope values based upon them, whereas values for trials 2, 3, 5, 6, 8, and 9 have played no role in determining the slope values reported above. As a means of eliminating this difficulty, and in turn to make more complete use of the available data, the slope analysis was redone on the basis of the best fitting logarithmic curves which could be fitted to the acquisition data for Conditions 0, 10, 0-cg, and 10-cg. A reasonably good fit to the empirical data was obtained with a logarithmic function. The results of this analysis are presented in Fig. 5.

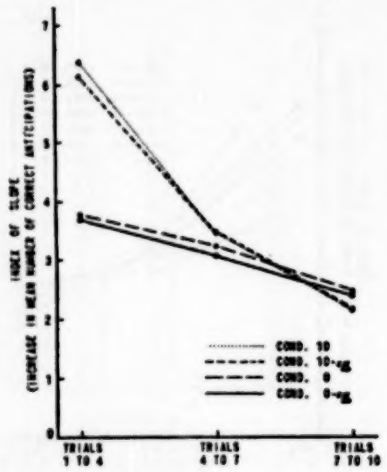

FIG. 5. Analysis of the slopes of learning curves for List B when slope is measured in terms of the increase in the mean number of correct anticipations from trials 1 to 4, 4 to 7, and 7 to 10. (Based upon fitted curves.)

The slope analysis data derived from the fitted curves confirms and reemphasizes the conclusion, suggested by the empirical data, that under the conditions of the present experiment the facilitating effects of paired-associate verbal learning and of non-associative color guessing upon immediately subsequent paired-associate learning act primarily during the first few trials in the learning of the new list.

There was relatively little difference between the various experimental conditions in terms of the rate with which learning progressed after the fourth trial in the learning of List B. However, it may be noted that the slope between trials 7 to 10 for those conditions receiving relatively large amounts of either preliminary verbal learning or color guessing practice tends to be somewhat smaller than that for those conditions which called for little or no preliminary practice. It appears that this tendency can be attributed to the fact that under the maximum preliminary practice conditions a moderate number of Ss achieved perfect or near perfect learning scores during the last few trials on List B and thus could show little or

no further improvement in learning. Under Conditions 10 and 10-cg, the mean number of correct anticipations on the tenth trial on List B was nearly 12 correct responses out of the 15-item list, whereas under Conditions 0, 0′, and 0-cg the mean was only 9.4 correct responses.

Discussion

Several possible hypotheses may be offered to account for the manner in which preliminary warm-up practice acts to increase the rate with which a new paired-associate list is learned after a brief interval.

One possible interpretation of the obtained results would be that the preliminary memory drum practice gave *S* a chance to 'catch his breath' and get both physiologically and psychologically calmed down before learning the test list. According to this hypothesis, the learner would be less subject to distraction from interfering or irrelevent stimuli after such a calming down period and therefore better able to learn the test list. The results of the present experiment do not support this hypothesis. Conditions 0 and 0′ from Experiment A called for identical experimental schedules on the first experimental day. On the second experimental day, *S*s under Condition 0 learned List B after engaging in 10 minutes of cartoon sorting, whereas individuals serving under Condition 0′ learned List B as their first and only experimental activity immediately upon reporting for their second session. If time to 'catch one's breath' is an important factor in influencing the rate with which a new paired-associate list is learned after a brief interval, then, to the extent to which a 10-min. cartoon sorting interval made this possible, *S*s serving under Condition 0 could be expected to learn List B more rapidly than did *S*s under Condition 0′. It has been shown that the difference in the rate with which List B was learned under these two conditions lacked significance when evaluated by the *t*-test.

It might be hypothesized that the facilitating effects of preliminary paired-associate practice were derived from the fact that something was *learned* during the warm-up practice, i.e., that the formation of paired associations during the preliminary practice was in some way responsible for the more rapid learning of the test list. In evaluating this hypothesis, it should be remembered that the present study showed that 10 warm-up trials on the non-associative color guessing task were essentially as effective in facilitating subsequent performance on List B as was a comparable amount (number of trials) of preliminary paired-associate learning practice on List A. It was also shown that the color guessing activity produced no general practice effect comparable to that produced by paired-associate learning practice. In the light of this evidence it would appear that *learning per se* is not an essential element in activities which have a facilitating warm-up effect upon subsequent paired-associate learning.

It is conceivable that preliminary practice on one paired-associate list produced its facilitating effect by providing *S* with a goal or standard with which he could compete during the learning of the test list. The fact that List B was learned at essentially equal rates under Conditions 10 and 10-cg is strong evidence against this interpretation. The color guessing task in Condition 10-cg could hardly be expected to have given as adequate a standard for self competition as did the paired-associate learning practice from Condition 10, and yet the differ-

ence between these two conditions on the learning of List B was negligible and lacked significance when tested by an analysis of covariance.

A fourth hypothesis to account for the manner in which preliminary memory drum practice acts to facilitate subsequent paired-associate learning after a brief interval stresses the role of response set and the conditions under which it is produced. Ammons (1) and Irion (2, 3) have both identified the warm-up effect with a loss of set to perform the activity involved. The warm-up period is thought to involve a reinstatement of the various postural and 'attentive' adjustments necessary for optimal performance. Both Ammons and Irion were dealing with the warm-up effect as it is found in relearning situations. Their formulation appears to be equally applicable to situations involving the learning of new materials, such as prevailed in the present study. It is suggested that preliminary warm-up practice enables *S*, at the time of the learning of the test list, to assume more rapidly the postural or 'attentive' adjustments necessary for optimal performance than they would have been assumed had no preliminary practice been received. In the present study, the stimuli which appeared in the window of the memory drum constituted only a part of the total stimulus complex to which *S* was responding. *S* was also responding to proprioceptive stimulation arising from his posture, muscular tensions, etc. It is thought that an important component of this stimulus complex arises from *S*'s adaptation to the rhythm of the exposure apparatus and from tensions developing out of his attempts to make correct anticipations whenever a stimulus item comes into view.

It appears reasonable to assume that *S*s who have had preliminary warm-up practice carry over some of the proprioceptive stimulation from the preliminary practice to the learning of the test list and are thus able to begin the learning of the test list at a higher level of learning efficiency than would otherwise be possible. Since the greater the amount of preliminary warm-up practice, the greater is the likelihood that *S* will have adopted the response set most appropriate for successful response, it follows that increased amounts of warm-up practice should be accompanied by correspondingly increased efficiency in the learning of the test list.

The results of the present study are in agreement with the above interpretation. In Experiment A it was found that the rate with which List B was learned increased directly with the number of trials received on List A 10 minutes previously. The fact that warm-up practice on the color guessing task and on List A were essentially equal in their facilitating effects upon subsequent performance on List B likewise supports the above hypothesis since the color guessing activity could be expected to produce essentially the same pattern of proprioceptive stimulation as would be produced by paired-associate learning practice.

Data from the present experiment do not provide an indication of the maximum limit to which the warm-up effect will rise as a function of increasing amounts of preliminary warm-up practice. As shown in Fig. 1, the magnitude of the warm-up effect bears an almost linear relation to the number of warm-up trials received over the range of from 0 to 10 warm-up trials which was investigated in the present study. Presumably, if extended so as to cover larger amounts

of warm-up practice, this function eventually would reach a maximum value.

Knowledge of the presence and operation of a warm-up factor in verbal learning situations makes it possible to interpret more readily the data reported above relative to the rate with which List A was learned under the various conditions in Experiment A. It will be remembered that in Experiment A, *Ss* received either 0, 2, 4, 6, 8, or 10 trials on List A on Day 1 and either 10, 8, 6, 4, 2, or 0 trials on List A on Day 2, the total number of trials on Day 1 plus Day 2 always equalling 10 trials. It was found that the mean number of correct anticipations on the total of 10 trials on List A was lower under each of the experimental conditions which called for a division of practice on List A between Day 1 and Day 2 than was the mean number of correct anticipations for *Ss* who learned List A either entirely on Day 1 or entirely on Day 2. This finding is not in keeping with the result which would have been predicted on the basis of most distribution of practice studies.

The operation of the warm-up factor provides at least a partial explanation for this apparent incongruity. It would appear that under those experimental conditions which called for a distribution of practice on List A over Day 1 and Day 2 the efficiency of performance on Day 2 was determined by at least two opposing factors: (1) the generally facilitating effects of distribution of practice, and (2) a retarding effect attributable to the need to warm-up. Under the conditions of the present experiment the retarding influence attributable to the need to warm-up appears to have more than offset any advantage which might otherwise have accrued from the distributed practice.

Summary

1. Two experiments were undertaken in an attempt to measure the effects upon the learning of a paired-associate list of varying amounts of immediately preceding memory drum activity when (1) amount of previous practice is held constant, (2) the effects of proactive inhibition are minimized, and (3) the interval between the warm-up practice and the learning of the test list is brief.

2. Experiment A involved 175 *Ss*, 25 serving in each of seven conditions. On the second experimental day the various conditions called for 10 trials on a test list of paired-associate nouns to be given 10 min. after completing either 0, 2, 4, 6, 8, or 10 preliminary warm-up trials on a paired-associate adjective list. The rate with which the test list was learned was found to bear a direct relation to the number of warm-up trials completed 10 min. previously.

3. Experiment B, which involved 50 *Ss*, 25 serving in each of two conditions, demonstrated that the rate with which a test list of paired-associate nouns is learned is facilitated by preliminary warm-up practice on a memory drum task entailing color guessing under conditions which do not involve associative learning.

4. The paired-associate learning practice from Experiment A and the non-associative color guessing task from Experiment B were approximately equal in effectiveness in facilitating the learning of the test list.

5. Analysis of the slopes of the learning curves for the test list from both experiments indicates that the facilitating effects of preliminary warm-up practice are most prominent during the first few learning trials on the test list.

6. The results of the present experiment can best be accounted for through the role of response set and the conditions under which it is produced. It is suggested that preliminary warm-up practice enables *S*, at the time of learning the test list, to assume more rapidly the postural and 'attentive' adjustments necessary for optimal performance than they would have been assumed had no preliminary warm-up practice been received. This transfer of a facilitating response set would permit *S* to begin the learning of the test list at a higher level of efficiency than would otherwise be possible.

(Manuscript received July 13, 1949)

References

1. Ammons, R. B. Acquisition of motor skill: I. Quantitative analysis and theoretical formulation. *Psychol. Rev.*, 1947, 54, 263–281.
2. Irion, A. L. Reminiscence and retention as a function of length of rest and amount of pre-rest practice. Ph.D. dissertation deposited in the State Univ. of Iowa Library, 1947.
3. Irion, A. L. The relation of "set" to retention. *Psychol. Rev.*, 1948, 55, 336–341.
4. Hamilton, C. E. The relationship between length of interval separating two learning tasks and performance on the second task. Manuscript in preparation for publication based upon a Ph.D. dissertation deposited in the State Univ. of Iowa Library, 1948.
5. Heron, W. T. Warming-up effect in learning nonsense syllables. *J. genet. Psychol.*, 1928, 35, 219–228.

STIMULUS PRE-DIFFERENTIATION AS A FACTOR IN TRANSFER OF TRAINING[1]

BY R. M. GAGNÉ[2] AND KATHERINE E. BAKER

Connecticut College

Introduction

One of the important characteristics of many motor skills is that their learning appears to depend to a greater extent upon a differentiation of stimulus aspects than it does upon an improvement in precision of muscular movement. In many motor tasks the individual may learn to respond to a very complex set of sensory cues with reactions such as pressing buttons, turning cranks, or pulling levers, which in themselves are easy and well practiced movements. Practical training on a skill of this sort often emphasizes the stimulus aspects of the task, though this emphasis may be given in quite an unsystematic manner. It is a common procedure, for example, to introduce a subject to the learning of a discriminative skill by giving him instructions whose main function is to point out the specific stimuli to which he must make the required reactions. In some cases a certain amount of practice in identifying these different stimuli, usually by means of verbal names, is also given before the learner undertakes to practice the motor task itself. The present experiment may be viewed as an investigation of the effectiveness for transfer of different amounts of practice of this preliminary sort, which presumably serves to make the stimuli more distinctive. The transfer of this type of training is measured in connection with a motor skill in which a number of different stimuli must be associated with the same number of different manual responses.

A number of previous experiments have been concerned with investigating the extent to which various kinds of preliminary training transfer to the learning of a motor task which was highly similar to the present one. In general these studies have interpreted the effect of preliminary training to be the reduction in the tendency of a learned response to generalize to stimuli which are similar to the one with which the given response has been associated. Thus, in one study (3), increasing positive transfer to a total skill was shown to follow increasing amounts of preliminary training on two of the four component habits of that skill. The amount of the resulting transfer was conceived to depend upon the degree to which conflicting generalized response tendencies to similar stimuli were reduced by preliminary training. In a second study (5), an analysis of the errors made in the final motor task following the same type of component training revealed that the reduction in errors was not confined to the specific S-R connections which had been practiced in the preliminary task, but applied also to S-R connections for stimuli which were only similar, in one or more aspects, to the ones used in preliminary training. Thus, it was demonstrated that discrimination training had a transfer effect which was not confined to specifically practiced responses to stimuli, but which could be interpreted as resulting from the reduction of generalization tendencies to all similar stimuli. In still a third experiment (4), the amount of

[1] This paper is a slightly modified version of Report 316-1-7 under Contract N7onr-316, Task Order I, between Special Devices Center, Office of Naval Research, and Connecticut College.

[2] Now at Perceptual and Motor Skills Research Laboratory, ATRC Human Resources Research Center, Lackland Air Force Base, Texas.

transfer to a motor task was related to the amount of preliminary training on a pictured representation of the task, in which the subject was required to make pencil marks in appropriate spaces, representing positions of switches to pictures of the stimuli actually used in the final task. Increasing positive transfer to the motor task was found to be correlated with increasing amounts of practice on the preliminary 'pictured' task. In this experiment, too, the results were in accord with the idea that preliminary training reduces interference within the task by bringing about a reduction in internal generalization.

In each of these cases, the effect of preliminary training may be said to be that of increasing the distinctiveness of the discriminative habits which must be learned in the final motor task. In this sense, the preliminary training has a 'pre-differentiating' effect which, in general, increases in degree as the amount of such training is increased. But in each one of these experiments some aspects of both stimuli and responses have been common to the preliminary and the final task, so that it is impossible to distinguish completely between 'stimulus pre-differentiation' and 'response pre-differentiation,' or even to determine whether such a distinction can legitimately be made. Thus, in the two studies in which preliminary training on components was given, subjects practiced two of the four responses (to two of the four stimuli) used later in the final task. In the experiment involving preliminary practice on a pictured representation of the motor task, although the responses were quite dissimilar to those of the final task, the directional aspects of the responses were the same in the two tasks (i.e., the position of the blanks to be checked by pencil corresponded in a relative sense to the position of switches to be pressed). While the effects of different amounts of discrimination training can be compared within each of these studies, in none of them was a clear distinction drawn between the effects of pre-differentiating stimuli and of pre-differentiating responses.

The present experiment was designed to measure the effectiveness of *stimulus* pre-differentiation in producing transfer to a motor task, under conditions in which the responses employed in the preliminary training were made to be as neutral and unrelated to those of the final task as possible. The subjects first learned to make verbal responses to a set of four light stimuli, differing from each other in color and position, and subsequently proceeded to learn a task which contained the same stimuli but required the responses of pressing four switches, one for each of the lights. Three different amounts of practice on the preliminary task were given to three different groups of subjects, in order to measure the way in which this variable affected transfer. The performance of each of these experimental groups was compared with that of a control group during 60 trials of practice on the final motor task. If stimulus pre-differentiation is a significant factor in the determination of transfer of training, it was expected that the performance of the experimental groups would be superior to that of the control group, which received no 'pre-differentiating' training.

The concept of stimulus pre-differentiation appears prominently in the set of hypotheses proposed by Gibson (6) to account for certain aspects of the learning of verbal paired-associates. These hypotheses, like the analysis of discriminative motor skills we have presented in previous papers, are based upon the idea that the acquisition of each required response generates tendencies to make the same response to stimuli similar to the stimulus with which the correct response is associated. Since the stimuli of the task are to some extent similar, these generalized response tendencies conflict with the other 'correct' response tendencies which are established to each

stimulus by direct practice. Continued practice may be expected to bring about a reduction in generalization, and consequently a reduction in the number of incorrect responses exhibited in the performance of the task. Some studies have found an initial increase in error occurrence to follow a small amount of practice on the task, and a subsequent decrease in errors as training is continued. This was the finding, for example, in studies of paired associate verbal learning by Gibson (7, 8) and by Gagné (1). Also, in the first of the present series of experiments (3), in which a 'paired-associate' motor task was employed, the occurrence of an initial rise in the number of overt errors at the beginning of the learning of the motor task was followed by a progressive decrease in errors as training continued.

The interpretation of these results leads to the hypothesis that any kind of preliminary training which serves to reduce the internal stimulus generalization of the task will have a facilitating effect upon the subsequent learning of the total task. Gibson's (6) proposition XV reads as follows: "*If differentiation has been set up within a list, less generalization will occur in learning a new list which includes the same stimulus items paired with different responses; and the trials required to learn the new list will tend to be reduced by reduction of the internal generalization.*" It is reasoned that if differentiation has been set up among a number of stimulus items it will be easier to differentiate them again later even though they are paired with different responses. This effect comes about because the preliminary learning has effected a reduction in the generalized response tendencies which interfere with the correct response tendencies to be established by practice. An experiment by the same author (8) showed that the number of overt errors in a second list of paired associates was smaller than the number occurring in a first list using the identical stimulus members but different responses.

With this introduction, the present experiment can be set in its proper theoretical framework. The attempt was made to reduce the internal generalization of a 'paired-associate' type of motor task by giving preliminary training on a task which utilized the same stimulus elements but quite different responses. It was expected that a small amount of such training would serve to bring about an initial increase in the internal stimulus generalization followed by a progressive reduction in generalization with continued training. A small amount of preliminary training should consequently result in an increase in the occurrence of incorrect responses, at least during the early stages of the learning of the final motor task. As the degree of pre-differentiation is increased by increased amounts of preliminary practice, a progressive reduction in errors should occur in the learning of the final task. It is to be noted that the attempt was made in this experiment to limit the effects of preliminary training to a reduction in *stimulus generalization*, by utilizing responses which were as different as possible from those of the final task. Thus, the responses were ones which were expected neither to interfere with nor to constitute practice on the final task responses. It may be said that the subject first learned to differentiate the stimuli on the basis of non-meaningful verbal 'names,' and then proceeded to utilize these same stimuli in learning a task involving motor responses of which he had no previous knowledge. We were interested in discovering the effectiveness for transfer of this process of making only the stimuli more distinctive.

Method

Subjects.—The subjects used in the experiment were 150 male undergraduate college

TABLE I

SCORES ON MATCHING CRITERIA FOR THE FOUR GROUPS OF THE EXPERIMENT. $N = 32$ FOR EACH GROUP

Measure	Group			
	Control	8-trial	16-trial	32-trial
Substitution Test (number right)				
Mean	62.87	62.78	62.69	62.69
S.D.	8.73	9.49	8.43	8.64
Reaction Time (unit = 0.01 sec.)				
Mean	49.74	50.71	50.32	50.28
S.D.	6.23	6.12	5.56	5.18

students. Most of these men had taken a course in elementary psychology, and were aware that the experiment was concerned with learning. They had no knowledge of the nature of the task, however, before the experiment. These subjects were paid for their services on an hourly basis. The men were assigned in random fashion to four groups of approximately equal size, one of these groups being the control group, and the others the three experimental groups. After matching, each group consisted of 32 men and the data obtained for these subjects were used in the analysis.

The groups were matched on the basis of a learning score obtained from the administration of the full-page Woodworth-Wells substitution test with a time limit of 80 sec., and also on the basis of average reaction time scores obtained from 25 trials of reaction to a single green light of the apparatus by pressing a single switch given at the close of the experimental session. The multiple correlation between these two measures and the total time of all 60 learning trials in the control group was .35, and between the same measures and total errors in the control group, .23. The means and standard deviations of the scores on matching criteria are given in Table I. Differences between these means are insignificant and in general indicate that matching of the groups was successfully carried out.

Apparatus.—The apparatus employed in the experiment was a modified form of one used in a previous experiment (3). A vertical display panel presents to the subject either a red or a green light near the top of the panel, and lights of the same color in a position near the bottom of the panel. The light stimuli are separated from each other by 7½ in. in the vertical dimension. A reaction panel which lies flat on a table contains a starting button in front on which the subject's thumb rests before each reaction, and four spring-return toggle switches, two on the left and two on the right of center. The switches within a pair are numbered 1 and 2, from left to right. The motor task requires the subject to learn to press one of the switches to each of the four different stimuli presented. The responses to be learned are as follows: Bottom red light—left switch No. 1; bottom green light—left switch No. 2; top red light—right switch No. 1; top green light—right switch No. 2. The subject is required to begin his response from a central starting button, and to respond as quickly as possible by pressing the correct switch as soon as a light appears on the display panel. An electric timer continues to run until the correct switch is pressed. The experimenter records both the time of reaction and the number and type of incorrect responses which are made.

One modification of the original apparatus makes possible the presentation of each pair of red and green lights in exactly the same position on the stimulus panel. Whereas in the apparatus previously used the two lights of each pair were placed one directly above the other, in the present apparatus both lights are behind the panel itself and illuminate a single ground-glass screen. This means that there is no longer any difference in position of the red and green lights of each pair. As has been stated, the pairs themselves are 7½ in. apart in the vertical dimension. A second modification of the apparatus was made to provide for the presentation of the letters which were used as responses in the preliminary task. In the center of the stimulus panel, between the two round openings for the lights, is a square aperture through which the response letters can be seen. Behind this square opening is mounted a shutter attached to a lever arm. This arm is tilted upward by means of a cam run by a synchronous motor mounted on the back of the panel. The motor is wired into the circuit which controls the operation of the lights, in such a way that it begins to run as soon as any light is turned on by the experimenter. After two sec., the cam pushes up the lever which operates the shutter,

exposing a letter to the subject's view. The letter remains exposed for a constant interval of approximately two sec., at which time the shutter closes, the motor stops automatically, and the stimulus light goes off. The letters themselves are mounted on a disc which can be rotated from the experimenter's panel by means of a flexible cable.

In the preliminary task the subject saw only the stimulus panel of the apparatus. The response panel with its starting button and reaction swtiches was completely covered by a removable wooden cover. Before each trial of the preliminary task, the experimenter first set the disc containing the four letters to a position in which the appropriate one of the letters could be exposed when the shutter opened. He next gave a ready signal and depressed a button which turned on the light corresponding to this letter. At the same time this button started the motor, so that after two sec. the shutter opened and exposed the letter which was to be associated with the light. The subject was to respond before the expiration of this two-sec. interval. The subject then saw the exposed letter for a constant interval of approximately two sec., after which the shutter closed and the apparatus automatically turned itself off.

Procedure.—At the beginning of each experimental session the full-page Woodworth-Wells substitution test was administered to each subject with a time limit of 80 sec. Following this, each subject in each experimental group was seated in a chair facing the apparatus and was given 8, 16, or 32 trials of preliminary training, depending on the particular group to which he was assigned. Subjects in the control group began the learning of the final task immediately. The preliminary task consisted of the learning of the following associations: Lower red light—'J'; lower green light—'V'; upper red light—'M'; upper green light—'S.' These particular letters were chosen to be as meaningless as possible with reference to the positions of the switches in the final motor task. Thus, such letters as 'L' which might have been associated with 'left,' or 'G' which might have been associated with 'green,' were avoided. The presentation of lights was in random order throughout each set of eight trials of preliminary training. Thus, the group given eight trials of preliminary training saw each stimulus light and its associated letter twice, the group given 16 trials saw each stimulus light and its associated letter four times, etc. The subject attempted to respond with the correct letter within two sec. after each light was turned on. Whether he responded or not at the end of two sec., the shutter opened and exposed the correct letter.

The instructions given to the subjects in each experimental group before beginning the practice of the preliminary task were as follows:

> On this panel you will see a red or a green light in this round hole (point), or a red or a green light in this hole (point). Only one light will appear at a time. Your task is to learn a letter which goes with each of these lights so that you can say it as quickly as possible after the light appears. The letters are J, V, M, and S. In just a minute I will tell you which letter goes with each light. As soon as you see one of the lights, you are to try to say the letter that goes with it as quickly as possible. After each light has been on for two seconds, the correct letter will appear in this slot so that you can see whether or not you gave the correct response. You will practice saying these letters for a number of trials. At the beginning of each trial I will say 'Ready,' before I press a button which turns on one of the lights.
>
> Now let me show you which letter goes with each light. First, you will see the light, and two seconds later the correct letter will appear in the slot. Try to remember them, so that you can say the letters fast and correctly on subsequent trials.

The four stimulus lights together with their associated letters were presented to the subject. The experimenter said the letters out loud as they appeared and he was encouraged to say them too.

> Now we're ready for the practice trials. Remember, try to learn to say the correct letter as quickly as you can after seeing the light, *before* it appears in the slot. Ready?

Following the administration of the proper number of trials of preliminary training, the wooden cover was removed from the response panel, exposing the switches to which reactions were to be made during the final task. The experimenter then proceeded to give the instructions for the final task, which were essentially the same as those used in a previous study (3). The subject was told to make as rapid a reaction as possible. It was explained that the bottom lights were to be associated with switches on the left, the top lights with switches on the right; and that a red light went with switch number 1, a green light with switch number 2 in each case. The control group, which did not practice the preliminary task, was also given the same instructions before beginning the learning of the motor task. For all groups, four different random orders of 60 presentations of the lights were employed in the final task, and equal numbers of subjects within each group had the lights presented to them in each order.

The time interval between trials was maintained at the constant value of approximately 10 sec. in both preliminary and final training of all groups. The time elapsing between the end of preliminary and the beginning of practice on the final motor task was two min.

At the end of 60 practice trials on the motor task, the reaction time of each subject was measured throughout 25 trials requiring responses to switch No. 2 on the left on the appearance of the lower green light. The average of these values was used in the matching of groups.

Results

Time of response.—Table II presents the average response times in successive sets of 10 trials for each of the groups of the experiment. The table also includes the standard deviations of these mean values. These data are shown graphically in Fig. 1.

It can be seen that, in general, the learning of the motor task proceeds in a negatively accelerated manner throughout its course. The learning of the control group begins at a relatively high time score value which is followed by a sharp reduction on the second set of 10 trials. Continued practice brings about additional improvement in performance throughout 60 trials. The experimental groups which received 8 and 16 trials of pre-differentiating practice on the preliminary task begin the learning of the motor task with a lower time score value than that of the control group. Some improvement in rate of learning as a result of these amounts of preliminary training is shown on subsequent trials, particularly during the middle stages of the 60-trial practice period. On the whole, though, the amount of this increase is not remarkably great. The experimental group which received 32 trials of preliminary training begins the learning of the motor task at a still lower time score value than those of the other experimental groups. Furthermore,

TABLE II

Average Time of Response per Trial in Successive 10-Trial Stages of Learning the Motor Skill, Together with the Standard Deviations of These Values, for Each of the Groups of the Experiment

Group	*N* of Scores	Trials					
		1–10	11–20	21–30	31–40	41–50	51–60
Control	320						
Mean		127.45	108.97	104.17	106.67	95.39	97.85
S.D.		67.92	55.91	54.23	57.04	44.52	44.25
8-Trial	320						
Mean		116.88	113.72	97.34	101.26	93.19	93.08
S.D.		61.35	66.96	47.83	55.73	47.37	46.97
16-Trial	320						
Mean		118.08	111.86	96.80	102.14	103.37	95.13
S.D.		60.82	65.26	51.07	53.50	44.63	50.07
32-Trial	320						
Mean		114.16	96.53	94.73	97.93	90.06	90.64
S.D.		29.85	45.90	49.03	51.58	39.08	42.57
t-values:							
Control—8-Trial		2.13*	1.72	3.41*	1.77	0.74	2.58*
Control—16-Trial		1.89*	1.24	1.90*	1.68	1.76	2.76*
Control—32-Trial		1.99*	3.98*	3.34*	2.16*	1.96*	2.80*

* The *t*-values are given for the differences between average values for control and experimental groups on each of the ten trials at each learning stage. Starred values are significant at the .05 level or better.

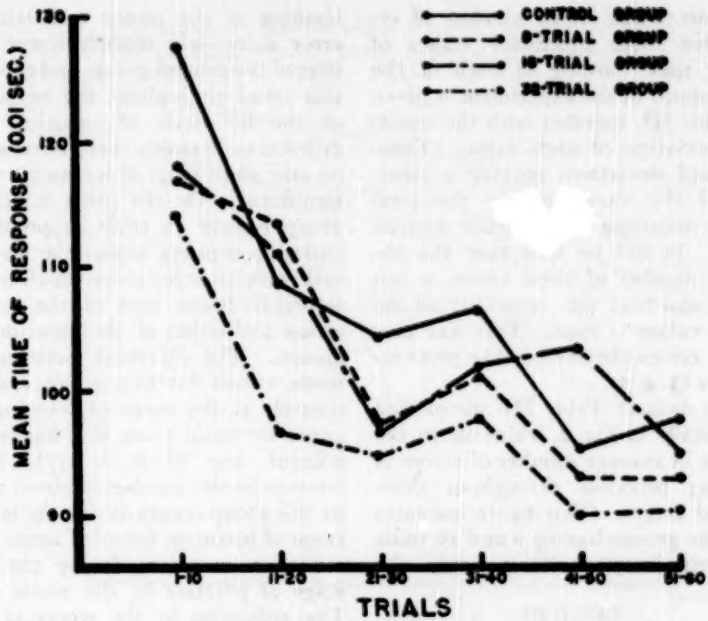

FIG. 1. Mean response time on successive 10-trial stages of learning the final motor task for groups having 0, 8, 16, and 32 trials of preliminary pre-differentiation training

this group continues to exhibit relatively large amounts of improvement in performance, compared with that of the control group, continuing until the final stages of practice have been reached. Thus, it would seem that 8 and 16 trials of pre-differentiating practice have resulted in a small degree of positive transfer to the motor task, and that 32 trials of such practice has brought about a considerably greater degree of transfer. Taking a value of 'total possible improvement' as the difference between the initial score of the control group and the best final score made by any of the experimental groups (cf. 2, Formula 6) it may be seen that the 32-trial group had accomplished 84 percent of the 'total possible improvement' by the end of 20 trials on the motor task, while the control group reached a learning level of only 50 percent after the same amount of practice on the motor task alone.

Table II presents the *t*-values of the differences between the mean time scores of the control and experimental groups during each of the six successive stages of learning. These values have been obtained by computing the standard errors of the differences between the means of each of the 10 trials within each stage, using $N = 10$. Values which are starred indicate significant differences at the .05 level or better in the direction shown by the data. It will be noted that the differences in scores between the control group and the group having 32 trials of preliminary training are significant at this level throughout the course of 60 trials. About half of those of the other two experimental groups also attain this level of significance.

Errors.—The mean number of errors for three successive stages of motor task learning in each of the four groups of the experiment is given in Table III, together with the standard deviation of each value. These standard deviations provide a measure of the variability in the total errors occurring during each 20-trial stage. It will be seen that the absolute number of these errors is not large, and that the variability of the mean values is high. This was true of the errors obtained in the previous studies (3, 4, 5).

The data of Table III are plotted graphically in Fig. 2, which shows the change in average number of errors as learning proceeds throughout three 20-trial stages. This figure indicates that the groups having 8 and 16 trials of preliminary training begin the learning of the motor task with an error score only slightly lower than that of the control group, and continue this trend throughout the remainder of the 60 trials of practice. The differences between these error scores on any given stage of learning are not significant. On the other hand, the group having 32 trials of pre-differentiation training begins the learning task with an error score which is considerably below that of the control group and either of the experimental groups. The difference between the mean values for this group and the control, at the stage of learning denoted by trials 1–20, is a highly significant one (C.R. = 3.57). Some increase in the number of errors made by this group occurs during the middle stage of learning, followed again by a reduction in errors during the final stage of practice on the motor task. The reduction in the scores of this group from those of the control group at the stage of learning denoted by trials 41–60 is significant at the five percent level (C.R. = 1.74).

TABLE III

AVERAGE NUMBER OF ERRORS IN SUCCESSIVE 20-TRIAL STAGES OF LEARNING THE MOTOR SKILL, TOGETHER WITH THE STANDARD DEVIATIONS OF THESE VALUES. CRITICAL RATIOS OF THE DIFFERENCES BETWEEN MEANS OF THE CONTROL AND EXPERIMENTAL GROUPS

Group	*N*	Trials		
		1–20	21–40	41–60
Control	32			
Mean		7.25	6.81	5.69
S.D.		4.65	4.30	4.31
8-Trial	32			
Mean		6.69	6.31	5.38
S.D.		5.92	4.73	4.43
16-Trial	32			
Mean		6.72	5.91	5.22
S.D.		4.94	4.71	3.18
32-Trial	32			
Mean		3.75	5.25	4.00
S.D.		3.16	3.73	3.46
Critical Ratios:				
Control—8-Trial		0.42	0.45	0.29
Control—16-Trial		0.45	0.80	0.50
Control—32-Trial		3.57	1.56	1.74

Fig. 3 shows the effectiveness of different amounts of pre-differentiating training in reducing errors made in the beginning stages of learning the motor task. The transfer effects of such training may be seen by comparing the average number of errors made by each of the groups on the first 20 trials. It is evident that the trend is the same as that previously described: a small, insignificant reduction in errors for the groups given 8 and 16 trials of pre-differentiation, and a considerable reduction in errors, as compared with those of the control group, for the group given 32 trials of pre-differentiating training.

Fig. 3 also shows a comparison of response time scores with the reduction in errors brought about by the different amounts of preliminary prac-

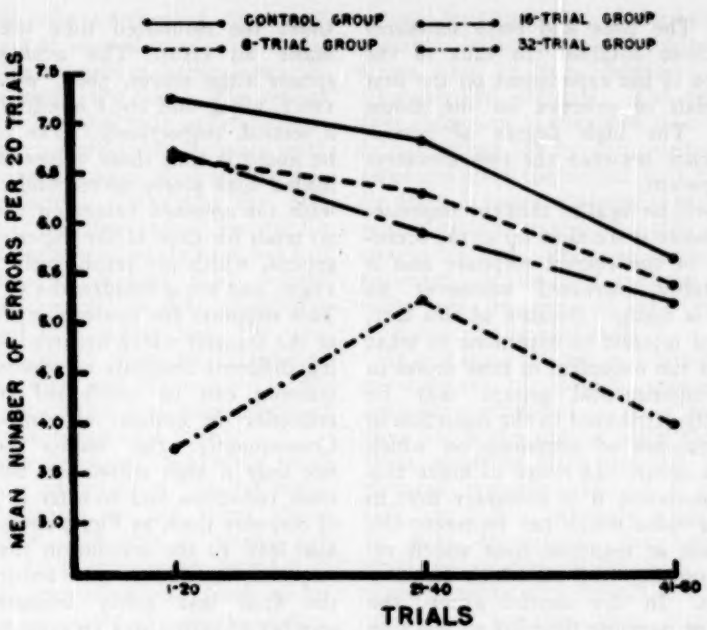

FIG. 2. Average number of errors made during successive 20-trial stages of learning the final motor task for groups having 0, 8, 16, and 32 trials of preliminary pre-differentiation training

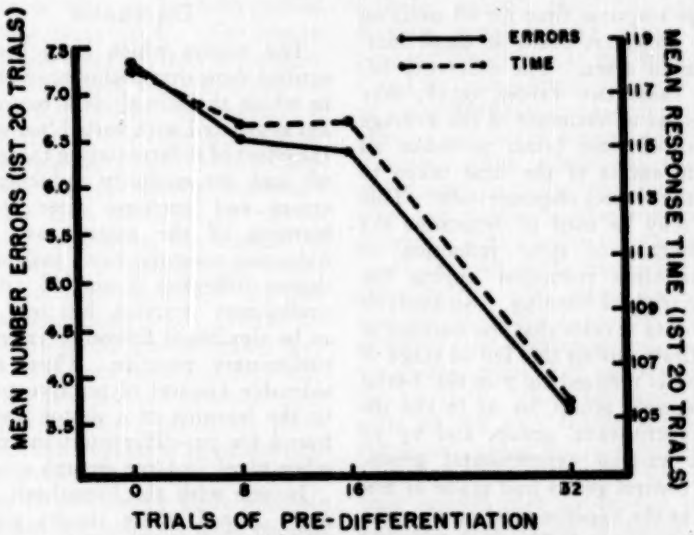

FIG. 3. The relation between number of trials of pre-differentiation training and two measures of performance on the first 20 trials of practice on the final task: (1) mean number of errors per 20 trials, and (2) mean response time per trial

tice. The time and error measures are those obtained for each of the groups of the experiment on the first 20 trials of practice on the motor task. The high degree of correspondence between the two measures is apparent.

It will be recalled that the response-time score is the time up to the occurrence of the correct response and is inevitably increased whenever an error is made. Because of this fact, it is of interest to determine to what extent the reduction in time scores in the experimental groups may be directly attributed to the reduction in the number of occasions on which errors occur. In order to make this determination it is necessary first to find a value which can represent the increase in response time which results on any trial on which an error occurs. In the control group, the average response time for all trials on which errors occurred is found to be 188.76 hundredths of a sec.; the average response time for all trials on which no errors occur is 90.58 hundredths of a sec. The difference between these two values, 99.18, may be used as an estimate of the average amount of time taken to make an error in excess of the time taken to make the correct response only. This value may be used to determine the contribution of error reduction to response-time reduction during the first 20 trials of learning. An analysis of our data reveals that the number of error trials during this initial stage of learning is reduced by 7 in the 8-trial experimental group, by 25 in the 16-trial experimental group, and by 75 in the 32-trial experimental group. If the control group had made as few errors as the experimental groups, the total response-time for 20 trials would be reduced by an amount equal to the difference in number of error-trials times the estimated time taken to make an error. The average response time scores, then, would be 116.2, 113.4, and 105.8 hundredths of a second, respectively. The fact to be noted is that these values of estimated time scores correspond closely with the *obtained* values on the first 20 trials for each of the experimental groups, which are respectively 115.3, 114.9, and 105.4 hundredths of a sec. This supports the contention that all of the transfer which occurred following different amounts of preliminary training can be attributed to the reduction in number of error-trials. Consequently, the results indicate not only a high correlation between error reduction and transfer in terms of response time, as Fig. 3 shows, but also lead to the conclusion that the response-time scores were reduced in the final task solely because the number of errors was reduced by the preliminary training.

Discussion

The results which have been described show that preliminary practice in which the stimuli of a motor task are associated with verbal 'names' has the effect of differentiating these stimuli and consequently reducing both errors and response time for the learning of the motor task. This reduction occurred to an insignificant degree following 8 and 16 trials of preliminary practice, but was found to be significant following 32 trials of preliminary practice. Thus, a considerable amount of positive transfer to the learning of a motor task was found for pre-differentiation training when given in large enough amounts.

In line with the hypothesis previously stated, these results may be viewed as the effect of stimulus pre-differentiation on the generalization tendencies within the task. Pre-

sumably, the number of 'incorrect' responses to stimuli provides a measure of the amount of internal generalization within the task. The results seem to indicate that generalization has been progressively reduced with increasing amounts of stimulus pre-differentiation. Evidently the procedure of learning verbal 'names' for each of the stimuli of the task has had the effect of making these stimuli more distinctive, so that when they are used in a motor task, the connections which must be established between them and the required motor responses exhibit a smaller degree of interference with each other.

It may be noted that while small amounts of stimulus pre-differentiation training do not bring about significant reductions in errors, neither do they produce an increase in internal generalization. Such an increase in internal generalization, shown by an initial increase in overt errors, has been found to follow small amounts of training in both verbal and motor tasks (3, 8). The reason that this phenomenon does not appear in this case cannot be determined directly from the data. However, it is possible that in the case of this relatively simple task, eight trials of pre-differentiation plus the amount which may have been provided by the instructions, represents an amount of training which is sufficient to reduce the internal generalization of the task from its maximum point. Thus, the point of maximum generalization may already have been passed in the learning of the group which received eight trials of preliminary practice. However, marked reduction in generalization from the amount existent without training requires more than 16 trials of practice, according to these results.

In general the effect of preliminary training has been to bring about decreased 'confusability' of the stimuli, thereby reducing the number of errors, and, in turn, decreasing the response-time measure. The finding that reduction in response time may be accounted for by error reduction alone confirms the notion that transfer depends upon reduction in internal generalization among the stimuli of the motor task, when, as in our experiment, preliminary training provides no practice on the motor responses of the final task.

It is of some interest to compare the effectiveness of *stimulus* pre-differentiation, produced in this experiment, with the sort of pre-differentiation resulting from practice on a pictured representation of a motor task, used in a previous study (4). In that experiment the subjects practiced a preliminary task composed of pictures of the actual stimuli employed, together with responses which were similar to the responses of the motor task in their directional aspects, though they consisted of making pencil marks rather than pressing switches. It was found that 8 and 16 trials of preliminary practice on the pictured representation brought about a significant amount of positive transfer, while additional trials of preliminary training produced somewhat less marked increases about this amount. The results of the present study contrast with these in that they show only a very small amount of transfer following 8 and 16 trials of stimulus pre-differentiation, and a far greater amount of transfer following 32 trials. Thus, although the results of the two studies cannot be compared quantitatively because of apparatus and subject differences they suggest that an amount of preliminary training of as much as 16 trials on the pictured representation is a much more effective kind of pre-

differentiation than is the strictly 'stimulus' pre-differentiation provided in the present experiment.

This comparison of results leads to the conclusion that something more than stimulus pre-differentiation must be involved in the kind of preliminary practice provided in a pictured representation of a motor task. It seems possible that the importance of providing preliminary practice on making responses in the correct directions was not sufficiently emphasized in the discussion of the previous experiment. When the stimuli are differentiated by being paired with verbal responses which do not include this directional aspect, as in the present study, more than 16 trials of preliminary practice appear to be required before significant positive transfer effects occur in the learning of the motor task.

It has been pointed out previously that the learning of motor skills in practical situations often includes preliminary practice, one function of which may be to differentiate the stimuli of the task. The results of this experiment indicate two things concerning the usefulness of such training. In the first place they show that the procedure of requiring the learner to identify the stimuli of a motor task by name does not have an adverse effect on the subsequent learning of the motor task. Secondly, these results show that if such pre-differentiating practice is carried on systematically throughout a number of trials, it may have a significant positive transfer value for the learning of the motor task. In practical terms, this implies that if a training aid is designed to depend for its effectiveness on the practice it provides in differentiating the stimuli of a motor task, a brief period of such practice is of little value. On the other hand, if a large enough amount of systematic practice in differentiating the stimuli is given, a considerable degree of positive transfer to the subsequent learning of the motor task may be expected.

Summary

The transfer of training effects of different amounts of preliminary training in 'stimulus pre-differentiation' were measured for a discriminative motor task. Four matched groups of 32 subjects were given respectively 0, 8, 16, and 32 trials of training in which they practiced the association of the letters 'J,' 'V,' 'S,' and 'M' to four different light stimuli on a panel. Following this, time and error scores were obtained for each subject during 60 trials of practice on a motor task in which the identical stimuli were paired with manual responses to four different switches.

1. The response-time scores in successive sets of 10 trials show that 8 and 16 trials of pre-differentiating practice bring about slightly faster learning of the motor task than shown by the control group (0 trials). However, marked facilitation of learning throughout the 60 trials occurs after 32 trails of preliminary training on stimulus pre-differentiation.

2. The error scores after 8 and 16 trials of pre-differentiating training is slightly and insignificantly reduced from those of the control group. A considerable, and significant, reduction in errors occurs in the group given 32 trials of preliminary practice, the greatest error-reduction occurring during the first 20 trials of learning.

3. The amount of transfer exhibited by the various experimental groups, measured in terms of response-time scores on the first 20 trials of learning, is closely correlated with the amount of error reduction on the same

20 trials. This result is consistent with the hypothesis that training on stimulus pre-differentiation brings about a reduction in the internal generalization of the task, thus reducing the number of incorrect responses and thereby bringing about facilitation of learning.

4. The results indicate that preliminary training which requires the subject to differentiate the stimuli of a motor task by learning their 'names' has a certain degree of effectiveness for transfer to the learning of the motor task. More than a small amount of such training is apparently required, however, in order for significant degrees of positive transfer to be obtained.

(Manuscript received August 8, 1949)

References

1. Gagné, R. M. The effect of sequence of presentation of similar items on the learning of paired associates. *J. exp. Psychol.*, 1950, **40**, 61–73.
2. Gagné, R. M., Foster, Harriet, & Crowley, Miriam E. The measurement of transfer of training. *Psychol. Bull.*, 1948, **45**, 97–130.
3. Gagné, R. M., & Foster, Harriet. Transfer of training from practice on components in a motor skill. *J. exp. Psychol.*, 1949, **39**, 47–68.
4. Gagné, R. M., & Foster, Harriet. Transfer to a motor skill from practice on a pictured representation. *J. exp. Psychol.*, 1949, **39**, 342–355.
5. Gagné, R. M., Baker, Katherine E., & Foster, Harriet. Transfer of discrimination training to a motor task. (To be published)
6. Gibson, E. J. A systematic application of the concepts of generalization and differentiation to verbal learning. *Psychol. Rev.*, 1940, **47**, 196–229.
7. Gibson, E. J. Retroactive inhibition as a function of degree of generalization between tasks. *J. exp. Psychol.*, 1941, **28**, 93–115.
8. Gibson, E. J. Intra-list generalization as a factor in verbal learning. *J. exp. Psychol.*, 1942, **30**, 185–200.

RETENTION AS A FUNCTION OF STAGE OF PRACTICE

BY RUTH GREENBERG AND BENTON J. UNDERWOOD

Northwestern University

Introduction

Requiring *S* to learn successive samples of a given kind of material may be thought of as a manipulation of stage of practice. It is well known that rate of learning is directly related to stage of practice, and, according to Ward's (5) data on the learning of successive lists of nonsense syllables, the relationship is a negatively accelerated one.

We have been unable to find a single report which has dealt with *retention* as a function of stage of practice. Is there a phenomenon of 'learning how to retain' that is analogous to 'learning how to learn'? Two considerations would indicate a positive answer:

1. In general, the faster the learning the better the retention. Therefore, if we observe a learning-how-to-learn effect we should expect also that retention would improve for each successive stage of practice.

2. If the basis for learning how to learn is to be sought in the improved adjustment of *S* to the learning situation, as suggested by McGeoch (1), we would anticipate that this improved adjustment would occur also for the retention tests.

In the present experiments we have used four stages of practice on paired-adjective lists and three retention intervals. *A stage of practice is defined as the learning, recall, and relearning of a given list.* For each stage of practice, 10 min., five hr., and 48 hr. were used as the three retention intervals.

Procedure

Three separate experiments are involved, each differing only in length of retention interval. The three experiments will be called the 10-Min. Conds., the 5-Hr. Conds., and the 48-Hr. Conds. In the 10-Min. Conds. *S* learned a list to eight out of ten correct responses, rested 10 min., and then recalled and relearned the list to one perfect trial. This constituted the first stage of practice. On the following day *S* went through exactly the same procedure, but with a different list, to complete the second stage of practice. On the two following days the procedure was repeated with other lists to provide four stages of practice.

In the 5-Hr. Conds. the procedure was exactly the same as for the 10-Min. Conds. except that five hr. elap. time between learning and recall of each successive list. In the 48-Hr. Conds. a single stage of practice required three days to complete. Since no *S* learned a new list on the same day that he recalled and relearned a previous list, the 48-Hr. Conds. required a longer period to complete than did the other two experiments. Actually, 13 days elapsed in the 48-Hr. Conds. from the learning of the first list to the relearning of the fourth list. The four stages were completed by learning the first list on Monday, recalling it on Wednesday; learning the second list on Thursday and recalling it on Saturday of one week, and then giving stages three and four on the same days of the following week.

Twenty-four *S*s served in each experiment. The experimental variable for each experiment was stage of practice; therefore, the four lists were counterbalanced so that each occurred equally often at each stage of practice. The results show that the four lists do not differ significantly in difficulty.

The 72 *S*s were secured from introductory psychology courses and were naive to verbal learning experiments. General instructions for learning were given on the first day and *S* was told that he would be expected to recall and relearn a list at a later time. At recall *S* was given opportunity to become fully prepared and was then asked to give as many responses as possible and continue with the list to mastery. In the 10-Min. Conds. *S* was kept in the laboratory and judged lengths of very short time intervals during the ten min. rest periods. In the 5-Hr. and 48-Hr. Conds. *S* was dismissed from the laboratory immediately after reaching the criterion of eight out of ten correct responses.

Each of the lists of 10 paired two-syllable adjectives was presented in three different serial orders by a Hull-type memory drum geared to the standard two-sec. rate. The four lists follow:

robust-heathen	vacant-narrow	oblong-tardy	lovely-resting
lucid-senseless	benign-senior	jagged-sublime	uphill-filmy
dismal-obese	discreet-brittle	fickle-upright	oldish-instant
center-naive	pensive-hardy	marine-ragged	pious-mature
vicious-unseen	minor-chosen	clouded-girlish	rancid-barbed
salted-aflame	winding-exempt	tiresome-icy	adrift-equal
yeasty-teeming	limber-jerky	hidden-ancient	drastic-latent
joyful-lengthy	recent-gallant	demure-folded	jellied-aware
feeble-valued	sallow-formal	witless-oaken	nimble-vagrant
nameless-bloated	fertile-obtuse	inert-noxious	human-waxen

Results

Original learning.—The mean number of trials required to learn the lists at the four successive stages of practice for the three experiments is shown in Table I. It will be remembered that the criterion of learning was eight out of ten correct responses. None of the differences among the means of the groups at comparable stages of practice is significant, although *S*s of the 10-Min. Conds. are somewhat more variable throughout than are those of the other two groups.

The most reliable estimate of the learning-how-to-learn curve for this

TABLE I

Mean Number of Trials to Learn at Each Stage of Practice

	1		2		3		4	
	M	σ_M	*M*	σ_M	*M*	σ_M	*M*	σ_M
10-Min. Exp.	12.38	1.27	10.25	1.10	9.88	1.04	9.42	1.11
5-Hr. Exp.	12.88	1.03	8.63	.62	9.17	.74	8.58	.79
48-Hr. Exp.	13.50	1.07	9.54	.70	9.71	.68	8.88	.93

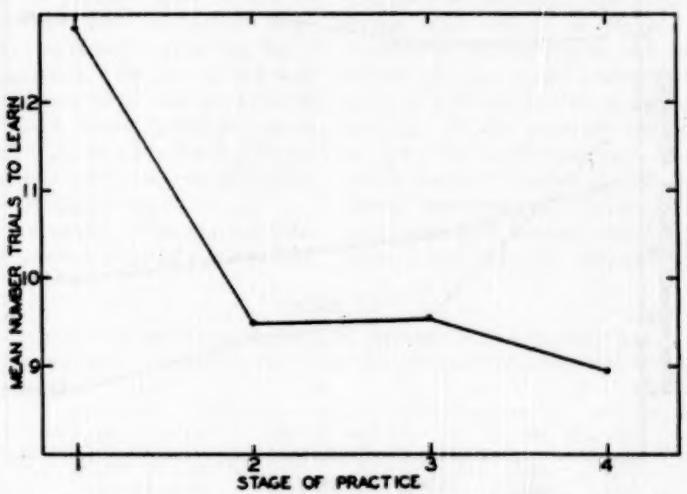

Fig. 1. Learning of paired-adjective lists as a function of stage of practice

TABLE II

MEAN NUMBER OF ITEMS CORRECTLY RECALLED AT EACH STAGE OF PRACTICE

	1		2		3		4	
	M	σ_M	*M*	σ_M	*M*	σ_M	*M*	σ_M
10-Min. Exp.	7.00	.30	6.92	.25	7.33	.21	7.25	.26
5-Hr. Exp.	4.92	.49	4.17	.44	3.88	.42	3.54	.48
48-Hr. Exp.	5.54	.36	3.04	.45	2.71	.51	1.96	.36

material is obtained by combining the data for the three experiments. The mean number of trials required for the 72 *Ss* to learn at each successive stage of practice is plotted in Fig. 1. It is apparent that the major amount of positive transfer takes place in a single stage of practice. Any further improvement is slight and statistically insignificant. Thus, with these lists of paired adjectives, the amount of learning-how-to-learn is much less than with serial lists of nonsense syllables.

Recall.—The mean number of items correctly recalled at each stage of practice is presented in Table II and graphed in Fig. 2. After 10 min., retention is unrelated to stage of practice. With a five-hr. retention interval recall is inversely related to stage of practice and this relationship is accentuated after 48 hr.

It will be noted that on the first stage of practice the recall is slightly better after 48 hr. than after five hr. Similar results have been found in other experiments (6) but the explanation for them remains obscure. The recall for the 48-Hr. Conds. drops markedly from the first to the second stage of practice and remains below that of the 5-Hr. Conds. for the remaining stages. The decrement in

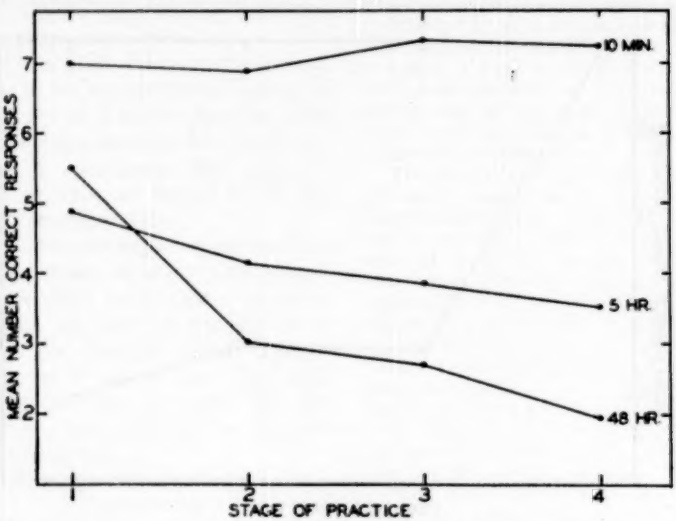

FIG. 2. Recall of paired-adjective lists as a function of retention interval and stage of practice

recall as a function of stage of practice is reliable for both the 5-Hr. and 48-Hr. Conds. In the 5-Hr. Conds. the difference between the recall from the first to the last stage is 1.38 ± .44, which gives a t of 3.14. This is significant beyond the one percent level of confidence. The difference is even greater for the 48-Hr. Conds., as is evident from the slopes of the curves in Fig. 2.

Relearning.—The mean number of trials required to relearn to one perfect trial for each stage of practice is shown in Table III. Two facts are to be noted in this table: (1) There is no apparent learning-how-to-*relearn* effect, and (2) the 48-Hr. *Ss* relearned somewhat faster on the average than did the 5-Hr. *Ss*. While the difference in relearning scores for the 5- and 48-Hr. Conds. is not reliable statistically, it is mentioned here as further evidence that five-hr. retention scores do not follow the projection of a smooth forgetting curve.

The relearning data make it clear that increased forgetting as a function of stage of practice is a transient phenomenon. This is further supported by an evaluation of the early relearning trials. On the second and third relearning trials the same trend (though much reduced) is shown as on the recall trial, but by the fourth relearning trial all evidences of forgetting have disappeared.

Errors at recall.—The inverse relationship between stage of practice and retention for the longer intervals, while unexpected, indicates the operation of proactive inhibition. In a previous study it was found that retention was an inverse function of the number of prior lists, being demonstrated when interference among lists was high and when a 25-min. retention interval was used (3). The present results suggest the same law for long time intervals and lists with low degrees of similarity. If proactive inhibition is to be invoked to account for the present results, the argument will be strengthened greatly if there is evidence for overt interference from previous lists at the time of recall of any given list. As shown in Table IV, such evidence is available. The 10-Min. Conds., which demonstrated no differential forgetting as a function of stage of practice, produced not a single intrusion from previous lists on any stage of practice. The 5-Hr. and 48-Hr. Conds., on the other hand, showed an appreciable number of intrusions from previous lists; enough to indicate that the response tendencies of the previous lists were interfering with recall. It is true that number of intrusions is not closely related to the actual amount of forgetting, but this is not an uncommon finding. In the previous experiment on proactive inhibition as a function of the number of prior lists (3) it was found that forgetting increased as the number of prior lists increased. However, there was no concomitant in-

TABLE III

MEAN NUMBER OF TRIALS REQUIRED TO RELEARN TO ONE PERFECT TRIAL

	1		2		3		4	
	M	σ_M	M	σ_M	M	σ_M	M	σ_M
10-Min. Exp.	4.17	.48	5.67	.68	4.17	.58	4.71	.56
5-Hr. Exp.	6.71	.76	7.88	.84	7.38	.65	6.96	.91
48-Hr. Exp.	6.67	.84	5.75	.50	6.38	.52	6.54	.76

TABLE IV

TOTAL INTRUSIONS FROM PREVIOUS LISTS AT THE TIME OF RECALL OF LISTS LEARNED AT STAGES 2, 3, AND 4

	2	3	4
10-Min. Exp.	0	0	0
5-Hr. Exp.	5	11	10
48-Hr. Exp.	5	5	9

crease in the number of intrusions. Thus, the intrusion data above can be submitted only as evidence that interference among the lists was taking place. The fact that the 5-Hr. Conds. showed a greater frequency of intrusions than did the 48-Hr. Conds. may be a function of the sample of *Ss* and not a function of the particular conditions. The 5-Hr. *Ss* showed more errors than did either of the other two groups during all learning sessions. For example, the erroneous responses made by the 5-Hr. *Ss* in learning all four lists totalled 978; for the 10-Min. Conds. the total was 859, and for the 48-Hr. Conds., 813. It is possible, therefore, that the greater number of intrusions on the 5-Hr. Conds. at recall may be attributed to characteristics of the *Ss* as such. Unfortunately, we have no knowledge of the relationships between *S* variables and intrusions.

DISCUSSION

The major empirical finding of the present experiments is that when the retention interval is from five to 48-hr. recall is inversely related to stage of practice. The general nature of the relationship and the error analysis lead us to interpret the finding as a demonstration of proactive inhibition as a function of time and number of prior lists. Certain aspects of this situation require consideration.

We must first inquire into the characteristics of the lists or conditions which produce the interference leading to proactive inhibition. The lists were constructed so that inter-list similarity would be low, although it is probably impossible to remove all similarities from among 40 adjectives. An inspection of the lists will show that there are at least remote similarities of one kind or another among some items. For most of the intrusions at recall it was impossible to discover the similarity factor operating. In a few cases, formal similarity seemed to have produced intrusions; in other cases serial position may have accounted for them.

There is, however, a common feature for all the lists, namely, the experimental situation, and this common factor may produce interference among the response systems when the response tendency for a specific stimulus is weakened by the usual forgetting process.[1] This common feature, plus some specific similarities among the items, may account for the interference producing the proactive inhibition.

It has been shown elsewhere (3, 4) that, on a logical basis, proactive inhibition must increase as a function of the time between learning and recall. A previous study failed to confirm this expectation when the two intervals tested were 20 and 75 min. The present study, according to our analysis, does confirm this expectation. After 10 min. there is no reduc-

[1] Miller (2) has suggested that early in the learning of a list the responses become cued to the gross context of the experimental situation, and only after several trials do the responses become cued to a specific stimulus. Miller's hypothesis might be extended to account for the present results. It is conceivable that with forgetting which takes place with the longer retention intervals, the responses remain cued to the gross context only. If this is the case, we would expect the interference among the lists in the present experiments because all are cued to the same context.

tion in retention as the number of prior lists increases; with five hr. there is some decrement and with 48 hr. the inhibition becomes marked. The longer the interval the less the discrimination among response systems.

The present study has one important methodological implication. In many studies of retention of verbal materials, counterbalancing or systematic randomization of conditions is employed so that each *S* serves in all conditions. In light of the present results it is apparent that stage of practice becomes an important source of variance when long retention intervals are used. In experiments where all *S*s are used in several conditions the response measures for a given condition are obtained at all stages of practice as a consequence of the counterbalancing, and, as we have shown, recall performance varies markedly as a function of stage of practice.

Summary

Three experiments were performed to determine the influence of stage of practice on retention. In one experiment the retention interval was 10 min., in another five hr., and in the third, 48 hr. A stage of practice consisted in learning, recalling, and relearning a list. Four such stages of practice were used in each experiment, hence four different paired adjective lists were required. The four lists were used equally often at each stage of practice and 24 *S*s served in each experiment. The results show:

1. Positive transfer in learning successive samples of the same class of material (learning how to learn) was largely complete after learning and relearning a single list. This result contrasts with the usual finding on serial lists of nonsense syllables where the practice effects are much more extensive.

2. With the 10-min. retention interval, stage of practice was unrelated to retention as measured by recall and relearning. With a five-hr. retention interval there was an inverse relationship between stage of practice and recall and no relationship with relearning measures. With 48 hr. between learning and recall the inverse relationship between stage of practice and recall was more accentuated than for the five-hr. interval, but again, no relationship was found by relearning measures.

3. An analysis of the intrusions at recall plus the general nature of the relationship between stage of practice and retention led to the conclusion that the findings represent proactive inhibition as a function of number of prior lists and time between learning and retention.

(Manuscript received July 21, 1949)

References

1. McGeoch, J. A. *The psychology of human learning.* New York: Longmans, Green, 1942.
2. Miller, R. B. Context and stimulus patterning in learning. Ph.D. dissertation, Yale Univ., 1947.
3. Underwood, B. J. The effect of successive interpolations on retroactive and proactive inhibition. *Psychol. Monogr.*, 1945, **59**, No. 273.
4. Underwood, B. J. Proactive inhibition as a function of time and degree of prior learning. *J. exp. Psychol.*, 1949, **39**, 24–34.
5. Ward, L. B. Reminiscence and rote learning. *Psychol. Monogr.*, 1937, **49**, No. 220.
6. Youtz, Adella C. An experimental evaluation of Jost's laws. *Psychol. Monogr.*, 1941, **53**, No. 238.

A QUANTITATIVE COMPARISON OF THE DISCRIMINATIVE AND REINFORCING FUNCTIONS OF A STIMULUS [1]

BY JAMES A. DINSMOOR

Columbia University

Introduction

Among the functional relationships which must presumably be included in a systematic description of behavior are those between the stimuli which are present in a situation and the strength or probability of a given response. In *The Behavior of Organisms* (11), Skinner uses the rate of bar-pressing by the white rat as a measure of these functions. If the cumulation of responses is recorded as a function of time, variations in rate are described by variations in the slope of the curve in a graphic representation or by the first derivative of the equation for this curve. It may be added that this derivative in turn represents the probability density of the response as a function of time.

The key concept in the conditioning of such operant or instrumental responses as bar-pressing is that of reinforcement. The defining observations for a reinforcing stimulus are that (1) responses which produce the stimulus in question increase in their rate of emission, and that (2) when the stimulus is later withheld, the rate of responding approaches but does not fall below the unreinforced or operant rate previously observed; that is, a 'reserve' of responses is established.

The possible roles of a stimulus are classified by Skinner in four categories: eliciting, reinforcing, discriminative, and emotional. The two roles which are to be compared in the present study are those of discrimination and what is known as conditioned or secondary reinforcement.

Discriminative stimuli.—The operations for the establishment of a pair of discriminative stimuli are the reinforcement of responding in one stimulus situation and the withholding of reinforcement in a differing stimulus situation. Since these two general situations must differ in one or more physical dimensions, two points on such a dimension may be abstracted and defined in terms of their selective relationship to the occurrence of reinforcement. The specific stimulus which has set the occasion for the reinforcement of a response may be labelled the S^D and that which has set the occasion for non-reinforcement of this response may be labelled the S^Δ.

The concrete training procedure followed by Skinner in the establishment of a discrimination involved the presentation of the S^D at a constant temporal interval (i.e., every five min.). The first depression of the bar which occurred in its presence produced the reinforcing stimulus and terminated the S^D. The alternative condition S^Δ continued until the next presentation of S^D; during this time, no further responses were reinforced.

The discriminative effects of a stimulus pair established in this fashion may be diminished or abolished by either of two techniques: responding may be reinforced in S^Δ or reinforcement may be withheld for all responses emitted in the presence of the S^D.

[1] Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Faculty of Pure Science, Columbia University.

The author is indebted to a number of colleagues for their theoretical and experimental suggestions and in particular to Professors William Nathan Schoenfeld and Fred S. Keller.

Two measures have been used for the difference in behavior which corresponds to the distinction between S^D and S^Δ. As in other functions, the standard measure is the rate of responding, which may be used for S^Δ during the training period or for either stimulus during a test period of continuous extinction. When only one response is made, however, as under the S^D in the standard training, or when the rate is contaminated by behavior incidental to the reinforcing operation (i.e., eating of pellets), this measure is no longer useful. In such cases, the period of time between the presentation of the stimulus and the occurrence of the first response, commonly known as the 'latency,' is sometimes substituted, although this procedure does not yield stable values (6, 11). However, the total number of responses obtained during a prolonged period of extinction is not a suitable measure in Skinner's opinion. Indeed, he is inclined to believe that the two stimuli act in a manner analogous to a filter, that they control the initial rate of emission without affecting the asymptote which the cumulative curve of responses will approach under either stimulus condition (11, p. 229).

Secondary reinforcing stimuli.—The concept of secondary reinforcement arises from the observation that an initial member may be added to a chain of responses which is now or was once followed ultimately by primary or unconditioned reinforcement. This initial member increases temporarily in its rate of emission, even though the total chain may no longer be followed by the original reinforcing stimulus. Such an effect is attributed to the action of a stimulus produced by the response in question, a stimulus which had no reinforcing power prior to the original training. For this stimulus, discrimination training is a sufficient establishing operation (11, pp. 246 ff.) and may be a necessary establishing operation (9, 10).

The operation which exhausts or extinguishes the reinforcing effect of such a stimulus is at least similar to that for the abolition of an S^D. It is the use of the stimulus, without further primary reinforcement, as itself a reinforcing agent for some response in the usual operant or instrumental procedure. That is, the stimulus is repeatedly produced by the response.

The effect of a stimulus, so utilized, on the response in question, is said to be reinforcing. In a test situation, the rate of emission is observed to increase, whether the response has previously produced primary reinforcement (1) or not (11, pp. 82 ff.), and the term 'reinforcing' implies that the total number of responses to be expected over an indefinitely prolonged period of extinction (the 'reserve') is likewise increased.

Present study.—The present study is designed to determine whether the two stimulus functions are equal or unequal in their effect on the rate of response when the governing variables are manipulated in the same fashion. All animals are first given extended discrimination training on the bar-pressing response, as an establishing operation. They are then divided into three groups, which are extinguished under corresponding stimulus schedules: (1) continuous presentation of the S^D until the response is emitted; (2) production of the S^D by the response; (3) complete absence of the S^D. The first procedure is considered a test of the discriminative function, the second a test of the secondary reinforcing function and the third an experimental control. Following three daily sessions of this extinction, the discriminative and secondary reinforcing groups are further sub-divided, and half of each group is transferred to the alternative stimulus schedule for the three remaining sessions.

If the cumulative curves of responding are recorded for the first three days of extinction, their height and shape may be compared. This comparison shows whether a differ-

ence in effect is obtained between the two main stimulus schedules, following identical training and during equal periods of unreinforced responding. Another comparison is obtained from the last three sessions of extinction, following the reversal of half of each experimental group. The rate obtained under either schedule of stimulus administration is tested following an equal period of extinction under the same usage or under the alternative usage. This shows whether the discriminative use of the stimulus subtracts to an equal degree from the consequent rate of responding under a reinforcing schedule, and whether the reinforcing use of the stimulus subtracts to an equal degree from the consequent rate of responding under a discriminative schedule.

Subjects and Apparatus

The organisms used in the present experiment were 36 male rats of albino Wistar stock, ranging in age from 90 to 100 days on arrival. Throughout the course of the investigation they were housed under environmental conditions which were permitted to vary as little as possible. The housing and experimental room was insulated from external sources of sound by a double door, and an electric fan was kept in continuous operation as an auditory masking device. Illumination was low, consisting of daylight coming through and around a blackout curtain and reflected light from a 25-watt lamp, which was shielded by a metal shade and a wooden partition from the table on which the living cages were kept. Temperature, with but occasional exceptions due to outside variation, was maintained from 78 to 80 degrees Fahrenheit.

Cages.—The living and experimental cages in which the animals were continuously housed throughout the period of experimentation were of galvanized iron, 10½ by 8 in. in floor space and 8 in. in height. They differed only in minor details from those described by Frick, Schoenfeld, and Keller (7). One end included an attached water bottle, which was removed only for replenishment of the supply, and an opening for the insertion of a tray of powdered food. The other end included an H-shaped opening, with a horizontal slot three and a half inches above the floor of the cage for the insertion of the bar. This end also bore a square food tube which opened above a small metal tray or cup on the inside.

Experimental chamber.—During the experimental sessions, these living cages were placed inside a chamber constructed of ¾ in. board and lined with a ½ in. layer of sound-absorbing celotex. The internal dimensions were 24 by 24 in. in floor space, 20 in. in height. This chamber was sound resistant and light proof except during S^D periods during the discrimination training, when it was necessary to open a 12 by 12 in. door in the front in order to permit the manual delivery of food pellets.

Pellets.—The food pellets used as the unconditioned or primary reinforcement were molded from wet Purina chow and dried. They were cylindrical in form and averaged .07 gm. in weight. These were dropped manually from a constant height of approximately 5½ in. through a plastic extension of the metal tube.

Bar.—The bar was a modified version of that in standard use in the Columbia laboratories (7). A cross-bar of ¼ in. hardened aluminum rod extended four in. horizontally, parallel to the cage front and approximately ¾ in. inside. On this cross-bar and its supporting shafts was folded a thin sheet of zinc four by ¾ in. This envelopment was largely successful in eliminating such responses as gnawing and shaking of the bar with the teeth. The height of the bar was 3½ in.; when depressed ½ in., which required a force between five and six gm., it actuated a microswitch and completed the recording circuit.

Stimuli.—Two levels of illumination were used inside of the experimental chamber. One condition consisted ordinarily of total darkness, slightly modified when the door in the front of the chamber was opened to permit reinforcement. The alternative condition was the illumination furnished by two 40-watt bulbs, laterally centered about four in. from the side of the cage and operated on 60 volts A.C. The indirect illumination through openings in the front, back and top of the cage averaged about three foot-candles and the illumination on an internal test-plate approximated .15 foot-candles. The change of illumination was controlled by the experimenter during discrimination training, by means of a hand switch, and by an automatic timing apparatus during extinction.

Measurements of the temperature inside the closed chamber revealed that in the course of prolonged light stimulation there was a rise to a maximum of about 90 degrees.

Records.—The occurrence and duration of all stimulus changes and all bar-pressing responses were recorded automatically on a waxed tape which moved at a constant rate of one mm. per

sec. (7). This recording device, together with a mechanical counter for tallying responses and apparatus for the control of stimulation during extinction, was kept outside of the experimental room. The only sounding apparatus inside the room was a solenoid in circuit with the bar, which served to set back a timing arm during discrimination training.

Procedure

Daily schedule.—For three days before the conditioning session, and thereafter throughout the experiment, each animal was placed on a daily schedule of 23 hours of deprivation and one hour in which dry powdered food was available. All experimental sessions were conducted in the preceding hour. It was also the practice to remove each animal from the cage for two or three minutes of fondling or gentling at some time later in the day.

Conditioning.—On the day preceding conditioning, a supply of pellets such as were to be used for reinforcement was placed on top of the powdered food. At the beginning of the conditioning session itself, five pellets were first delivered simultaneously through the tube and their consumption observed. Ten pellets were then delivered in discrete succession, each arriving after the last previous was consumed, in order to further condition the approach to the tray and to reduce the latency of this approach following the auditory and visual stimuli which signalled the arrival of the pellet. Finally, the bar was inserted and 10 depressions were reinforced. Reinforcement was judged to have occurred when the animal was observed to find and commence eating the pellet within a few seconds after releasing the bar. There were relatively few occasions on which animals pressed without meeting this criterion.

Discrimination training.—The training procedure which was used to establish a stimulus control over responding was a modification of the original schedule followed by Skinner (11). The first response made in the presence of the S^D, which was light for one half of the animals and darkness for the other half, was reinforced by the delivery of a pellet of food. The criterion for termination of the S^D was made the seizure of the pellet and the beginning of eating, in order to avoid the possibility that the S^Δ for bar-pressing might become an S^D for approaching the tray. After the first 20 or 30 trials, however, the animal's reaction became faster than the experimenter's.

The duration of the alternative stimulus condition, S^Δ, was not determined by a constant interval, as in Skinner's work, but by the animal's behavior. The criterion for the termination of S^Δ was the lapse of a period of 30 sec. without a response.[3] This criterion had two purposes: (1) to prevent secondary reinforcement or delayed primary reinforcement of responses in S^Δ due to occasional appearance of the S^D immediately following a response; in no case was a response in S^Δ followed either by the S^D or by a pellet within 30 sec.; (2) to provide more time for the extinction of S^Δ responding to those animals characterized by relatively high rates, thus tending to reduce the variability in the later test period of extinction.

Ten cycles of discrimination training, involving 10 reinforcements, were given immediately after conditioning, as a part of the first experimental session. On the six succeeding days of experimentation, the following number of cycles were given: 10, 15, 25, 30, 50, and 60. Thus, each animal received a total of 200 presentations of S^D and S^Δ. This schedule was based upon exploratory data and was intended to limit the experimental session to no more than 50 min.

Extinction.—Six daily 50-min. sessions of extinction were given to each rat. Two stimulus schedules were utilized for the experimental groups: (1) during most of the session, the stimulus condition (which may be termed 'prevailing' stimulus) was light, but each response was followed by three sec. of darkness; (2) the prevailing stimulus was darkness, but each response was followed by three sec. of light. This three-sec. period was timed automatically from the beginning of each bar-depression and was re-set by release of the bar. Thus, a further response during the three-sec. period led to its extension. Since this dichotomy of schedules bisected the dichotomy formed by the use of light or darkness as the S^D in training, in a factorial design, two of the sub-groups thus formed were presented with the S^D until the occurrence of the response—a test of the discriminative role of the stimulus—and the other two sub-groups produced the S^D by responding—a test of the secondary reinforcing role. It should be noted that the stimulus conditions under which these two roles were tested are matched by this design.

Following the first three daily sessions of extinction, each 50 min. in duration, half of the animals on each schedule were transferred to the alternative schedule, and extinction was continued for three more daily sessions of the same duration. This additional dichotomy bisected the experimental groups once again. The 32 experimental animals were thus divided and sub-

[3] This type of criterion has been in standard use in undergraduate laboratories at Columbia University. It was first suggested by Mr. D. E. Page.

TABLE I

A Summary of the Experimental Design

For each sub-group of four rats the prevailing stimulus condition is listed for each stage in the procedure, the stimulus role which is tested, and the sequence from the first to the second half of extinction.

Group	Training		First 3 Days Extinction		Last 3 Days Extinction		Sequence
	Prevailing Stimulus	S^D	Prevailing Stimulus	Procedure	Prevailing Stimulus	Procedure	
A	dark	light	light	S^D	light	S^D	same
B	light	dark	dark	S^D	dark	S^D	same
C	light	dark	light	S^γ	light	S^γ	same
D	dark	light	dark	S^γ	dark	S^γ	same
E	dark	light	light	S^D	dark	S^γ	reversed
F	light	dark	dark	S^D	light	S^γ	reversed
G	light	dark	light	S^γ	dark	S^D	reversed
H	dark	light	dark	S^γ	light	S^D	reversed
J			same	S^Δ	same	S^Δ	

divided by three orthogonal criteria of classification into eight sub-groups.

For control purposes, four additional animals were extinguished in the absence of the S^D. In two cases, the stimulus condition was light, in two cases darkness. The complete design is summarized in Table I.

Statistical Treatment

The statistical techniques used for the evaluation of the null hypothesis in the present experiment were analysis of variance and analysis of covariance. The factorial design permitted the segregation of several variables contributing to the experimental effects. For example, such effects as those due to the prevalence of light or of darkness during extinction should not be included in the estimate of random variation derived from the variance among individual animals within their respective groups. This might lead to an exaggeration of the estimate and a consequent diminution in the sensitivity of the test. The data show that this treatment was well advised in the present instance.

The use of analysis of covariance is roughly equivalent to the procedure of preliminary matching of the animals in the experimental groups on a measure believed to be related to the dependent variable. Four difficulties in matching, however, were surmounted by this technique. First, the best predictor or set of predictors for the extinction totals was not known in advance. Second, the predicting variables actually used could not have been obtained prior to the assignment of animals to their respective groups. Third, matching could not be made as accurate as adjustment of the original values by means of a regression equation. And finally, no appropriate statistical treatment is at present available.

The choice of predictive variables for the analyses of covariance was made after the examination of a number of rank-difference correlations. In view of the small size of the coefficient in each case and the small variance between group means, it seemed unlikely that a substantial change in the *F*-ratios would be produced (8). In the light of negative findings, however, a failure to disprove the null hypothesis, it seemed desirable to make the test as sensitive as possible. The calculations followed the routine outlined by Fisher (5). Determination of the intraclass correlations incident to these analyses also made possible slight adjustments in the more important group means.

A characteristic difficulty encountered in the statistical treatment of operant responding is that the distributions of latencies, rates and extinction totals are limited at the lower end by the numerical value of zero but are not effectively limited at the upper end. Such a circumstance tends to produce skewed distributions, and this tendency was actually observed both in the training and the extinction data, although moderate in the latter case. Furthermore, occasional extreme values were obtained which seemed, by subjective observation, to be associated with behavioral artifacts. In the training period, for example, long latencies of response to the discriminative stimulus were produced by the emission of abortive responses, inadequate in force to close the recording circuit, and by drinking from the water bottle at the opposite end of the cage. In extinction, similarly, high rates were produced by the interven-

tion of such responses as gnawing, shaking the bar, jumping, climbing, poking the nose through the slot in the cage, and a 'trotting' motion of the forepaws across the width of the bar. In the absence of complete experimental control, it seemed desirable to minimize the contribution of these possibly spurious responses to measures of central tendency and scatter. For this reason, and to correct the mild positive skewing of the distribution, all analyses of variance and covariance were based upon the logarithms of the original cumulative response totals, and geometric means were used to express the relative performance of the several groups. An exception was necessary in the treatment of the latency and number of responses in S^{Δ} during discrimination training, due to the presence of zero readings; here medians were substituted for the respective geometric means.

Results and Discussion [3]

Training.—The first seven days of experimentation were devoted to conditioning and discrimination training. The animals received 200 presentations of the S^{D}, with consequent reinforcement in each case, and 200 presentations of S^{Δ}, under which reinforcement was withheld. As indicated in Table II, the number of responses declined in successive cycles of S^{Δ}, and the period of time between the onset of S^{D} and the occurrence of the response, often termed the 'latency,' became progressively shorter. The data indicate that the decrement in each case was most rapid in the early stages of training and that little further change took place in the later stages. This suggests that the difference in effect between the two stimulus conditions was close to the maximum which could be achieved by the present procedure.

An interesting phenomenon, which may be peculiar to some specific aspect of the present training procedure, was that in the last series of cycles the typical behavior consisted of the emission of a single response in the first few seconds of the S^{Δ} period.

In order to avoid confusion in discussion between the use of a stimulus as an S^{D} in training and the test of a discriminative schedule in extinction, the stimulus which functioned as S^{D} in training will hereafter be termed 'positive' or 'effective.'

First three days of extinction.—The initial test period consisted of three 50-min. sessions on three successive days. This period served to compare the rate of bar-pressing and the changes in rate when the effective stimulus was, first, presented prior to the response, as a discriminative stimulus, or, alternatively, was produced by the response, as a secondary reinforcer. The mean number of

[3] For tables giving the individual response totals and the raw data necessary for the analyses of covariance described below, order Document 2829 from American Documentation Institute, 1719 N Street, N. W., Washington 6, D.C., remitting $0.50 for microfilm (images one in. high on standard 35 mm. motion picture film) or $0.50 for photocopies (6 × 8 in.) readable without optical aid.

TABLE II

Separation in Strength of Reflexes during Discrimination Training

The median latency in seconds of the first response in S^{D} and the median number of responses emitted in S^{Δ} are presented for each tenth cycle.

Cycle No.	10	20	30	40	50	60	70	80	90	100
Responses	0.3	11.5	7.5	3.5	2.5	2.0	1.3	1.4	1.0	1.1
Latency	17.0	8.5	4.3	4.5	4.7	5.5	2.7	2.5	2.0	2.9
Cycle No.	110	120	130	140	150	160	170	180	190	200
Responses	1.0	0.9	0.9	0.9	1.3	0.6	0.8	0.8	0.8	0.8
Latency	2.5	2.5	3.2	2.8	1.5	2.0	1.7	2.8	1.9	2.0

bar-pressings for these two groups, and for the control group, which did not receive the effective stimulus at any time, are cumulated and plotted in Fig. 1 at five-min. intervals for the first day of extinction. It will be noted that the curves for the two experimental groups attain their maximum separation toward the end of this session.

The total responses cumulated for days one, two and three are represented by geometric means, and the final set of means is corrected by the use of a regression coefficient derived from the duration of the last 50 cycles of training for each animal. These are presented in Table III, along with the means of those animals which were extinguished in the light as a prevailing stimulus and those which were extinguished in the dark. Comparison of these means indicates that: (a) a substantial and progressive difference in totals is associated with the prevailing condition of illumination; (b) the rate of response is quite low, relatively, in the continued absence of the effective stimulus; and (c) the groups performing on discriminative and secondary reinforcing schedules achieve almost identical totals at the end of the second and third days.

TABLE III

GEOMETRIC MEANS FOR FIRST HALF OF EXTINCTION

Geometric means of the cumulative totals of responses at the ends of days one, two, and three of extinction, classified by prevailing stimulus condition and the role of the stimulus. The corrected value for day three is based upon the regression of the original totals with the duration of the last 50 cycles of training.

Classification	Geometric Means			
	1	2	3	3 (cor.)
Prevailing Stimulus				
Light	77.4	138.8	192.3	195.0
Darkness	121.2	213.2	287.8	283.9
Stimulus Role				
S^D	93.0	172.2	235.6	233.8
S^r	100.9	171.9	234.9	236.7
Control				
S^Δ	48.4	84.1	128.5	

The statistical significance of these differences in cumulative totals for days one and three was evaluated by analysis of covariance, utilizing the duration of the last 50 cycles of training as the correlated variable (.12 and .14). These analyses concurred in showing that the difference due to the

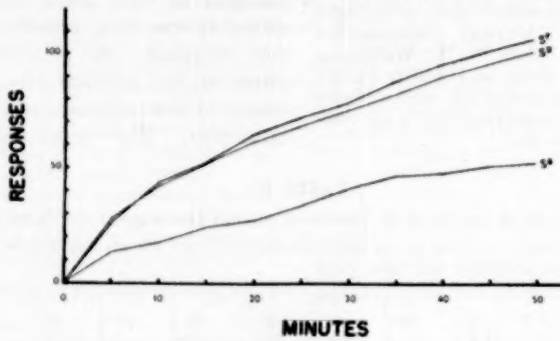

FIG. 1. The mean number of responses emitted by three groups of rats, cumulated at five-min. intervals during the first session of extinction. The first group extinguished in the presence of the stimulus which had previously set the occasion for reinforcement (S^D), the second received this stimulus for three sec. after each response (S^r), and the third or S^Δ group was extinguished in the absence of the stimulus, as an experimental control.

prevalence of light or darkness was significant, the chance probability in each case being less than one in a hundred. Whether this effect can be attributed to the photic or the thermal aspects of the stimulus, however, cannot be determined, for the two aspects have been confounded. On the other hand, the difference between the two stimulus functions, S^D and S^r, yields an F-ratio of less than one; the data afford no basis for rejection of the null hypothesis.

Contamination of the tests.—It may be objected that neither the test procedure for a discriminative role nor the test procedure for a reinforcing role is completely distinct operationally from the other. Each time that the positive or effective stimulus is produced by the response (reinforcing operation), it is possible for a subsequent bar-pressing response to follow in the presence of the stimulus (discriminative operation). And each time that the positive stimulus reappears following the response in its presence (discriminative operation), it may serve as a delayed secondary reinforcement for this response.

It is extremely improbable that the correspondence in height and shape of the cumulative response curves produced under either test schedule is a coincidence resulting from the chance balancing of two stimulus functions in each case. There are three systematic ways in which such a quantitative result could be produced, and each of these must be evaluated:

1. The correspondence of the curves might be produced by completely pure and distinct tests of two stimulus functions which were equal in their effects on responding. In the light of the foregoing discussion, however, such an assumption seems to be untenable.

2. The correspondence of the curves might be produced by the summation in each test of the two stimulus functions, if the discriminative effect of the stimulus were equally strong in both cases at all times during the test and if the reinforcing effect of the stimulus were equally strong in both cases at all times during the test. Neither of these equalities appears to be possible: a delayed reinforcement (during the discriminative test) does not maintain as high a rate of responding as an immediate reinforcement (11); a discriminative stimulus which is limited to three sec. in duration (test of secondary reinforcement) cannot be as effective in the production of a consequent response as a discriminative stimulus which is maintained until the occurrence of the response.

Furthermore, if the contamination is assumed to be complete in the tests used in the present study, operational separation of the two stimulus roles for measurement purposes would seem to be impossible. For example, if the bar were removed after each reinforcement with the positive stimulus, in order to preclude a response due to its discriminative role, rate of responding would no longer be available as a measure, and previous data shows that other measures such as latency and magnitude are not satisfactory (6, 11). Also, the removal of the bar would itself involve the removal of one of the stimuli which was present during the original establishment of the strength or effectiveness of the stimulus.

Similarly, if the discriminative stimulus were removed for a substantial period of time following each response, say 30 sec., rate of responding would again be lost as a measure, and the same difficulty follows.

Thus, the assumption of total contamination seems to be untenable. In any case, such an assumption would cast serious doubt on the possibility of operational separation of the two stimulus roles.

3. The assumption which is tenable is that some responses are produced discriminatively in the reinforcement test and that some responses are produced by secondary reinforcement in the discriminative test. This assumption is tenable if a second assumption be granted, that the discriminative use of the stimulus to produce a response exerts an equal weakening influence on the

reinforcing function of the stimulus, and that the reinforcing use of the stimulus exerts an equal weakening influence on the discriminative production of responses. According to these assumptions, there is a common 'reserve' of responses or resistance to extinction under the control of the positive stimulus, and this rate and total may be discharged either by presenting the stimulus discriminatively (prior to the response), by presenting it as a reinforcement (following the response), or by combining both methods of presentation. Supporting evidence for this contention is presented in the following section.

A rough indication of the actual degree to which the record of S^r responding is contaminated in the present study by the discriminative influence of the secondary reinforcing stimulus may be gained from a comparison of the temporal distribution of responding in the two experimental groups. Under the secondary reinforcing schedule, the effective stimulus is presented for three sec. following each response; the next succeeding response will often occur before the end of this period of stimulation. Similarly, under the discriminative schedule, the effective stimulus is removed for three sec. following each response; again, the next succeeding response may occur before the three sec. have elapsed. In the secondary reinforcing case, however, where the stimulus condition is positive, 41.4 percent of all responses during the first session of extinction are emitted under these circumstances, but in the discriminative case, where the stimulus condition is negative, only 28.8 percent of all responses are emitted during the same period. The difference of 12.6 percent, therefore, may presumably be attributed to the discriminative action of the positive stimulus under the secondary reinforcing procedure.

Last three days of extinction.—Following the initial test period, half of the discriminative animals were placed on a secondary reinforcing schedule and half of the secondarily reinforced animals were placed on a discriminative schedule. Extinction was then continued for three more daily sessions for all animals. The geometric means of the group performances on this test are presented in Table IV. The differences have been evaluated by an analysis of covariance, utilizing the cumulative response totals for the first three days of extinction as the correlated variable (.18). The difference in illumination is probably still a factor in the rate of response, since there is a sharp 'break' or discontinuity in the individual curves (not shown) and since the difference in means would be produced by chance less than five times in a hundred. The reversal of stimulus schedule, however, is not itself a significant factor; indeed, the animals which were reversed show somewhat less responding, on the average, than those which remained on the same schedule.

It may be concluded that unreinforced use of a stimulus for discriminative purposes produces an equal loss

TABLE IV

GEOMETRIC MEANS FOR LAST HALF OF EXTINCTION

Geometric means of the cumulative totals of responses from the beginning of day four to the ends of days four, five, and six of extinction, classified by prevailing stimulus condition and by the sequence in procedure from the first to the second half of extinction. The corrected value for day six is based upon the regression of the original totals with the total of responses during the first three days.

Classification	Geometric Means			
	4	5	6	6 (cor.)
Prevailing Stimulus				
Light	53.3	83.4	115.1	113.6
Darkness	71.7	135.4	184.6	187.0
Sequence				
Same	70.9	112.4	157.5	157.8
Reversed	53.9	100.5	134.9	134.6
Control				
S^{Δ}	24.9	49.0	67.8	

in its effectiveness for secondary reinforcing purposes and that unreinforced use of a stimulus for secondary reinforcing purposes produces an equal loss in its effectiveness for discriminative purposes. In short, the two stimulus uses may be interchanged as operations for diminishing or abolishing the effectiveness of the stimulus.

Theoretical Implications

The strength of members of a chain.— The present study constitutes an attempt to segregate so far as is operationally possible two stimulus-response sequences ($R \rightarrow S^r$ and $S^D \cdot R$) which would normally be observed as successive, or indeed overlapping, segments of a more extensive stimulus-response chain (. . . $R \rightarrow S^{r,D} \cdot R$ . . .). The original observations of chaining made by Skinner (11) were based upon a chain which included such items as bar-pressing, magazine sound, approach to the food tray, sight of pellet, seizure and so on. These were formulated in terms of the 'reflex' ($S^D \cdot R$), which he considered the element of the chain. Two basic relationships in strength between successive reflexes in a such a chain were noted:

First, it was found that an S^D which had set the occasion for the reinforcement of one response could be used as a reinforcing agent to add a prior member to the chain. That is, a response occurring at an operant or extinction rate could be increased in emission, both in terms of rate itself and in terms of resistance to extinction. Furthermore,

> The number of responses due to the reinforcing effect of the sound of the magazine is again of the same order as in the initial extinction curve itself. It is tempting to suggest the following law: a discriminative stimulus (such as the sound of the magazine) used as a reinforcement in the absence of ultimate reinforcement creates in another reflex a reserve just equal to that of the reflex to which it belongs (11, p. 105).

The logic of such a deduction from the data given is difficult to follow, but credit should be given for the formulation of a statement which is verified by the present study.

Skinner's second observation was that a chain of reflexes could be 'broken' at any arbitrary point by withholding one of the reinforcing stimuli in the sequence. For example, when the food pellet was withheld, both bar-pressing and approaches to the tray appeared to extinguish; or, when the sound of the food magazine was withheld, only the bar-pressing extinguished. This finding was formulated as the Law of Extinction of Chained Reflexes: "In a chain of reflexes not ultimately reinforced only the members actually elicited undergo extinction" (11, p. 105).

One amendment of this law seems to be necessary on the basis of previous data and one extension is indicated by the present findings. First, the correction stems from the observation that when one response (i.e., bar-pressing of a certain force or duration) is strengthened by reinforcement or weakened by non-reinforced emission, similar responses (i.e., of greater force or duration or of differing topography) seem to be strengthened or weakened simultaneously (2, 11, 13). "A dynamic change in the strength of a reflex may be accompanied by a similar but not so extensive change in a related reflex, where the relation is due to the possession of common properties of stimulus or response" (11, p. 32). No position will here be taken as to whether so-called 'response induction' is a direct consequence of the similarity of the responses themselves or a result of the similarity in the

proprioceptive stimulation which they produce; indeed, the distinction would be difficult to test. However, it is suggested that such a relationship must hold between successive members of a chain, and this factor must be kept in mind in any consideration of the extinction of chained reflexes.

Second, it will be noted that Skinner's law deals only with the intact reflex ($S^D \cdot R$), possibly due to a realization of the difficulty of using a stimulus as a secondary reinforcer without at the same time causing the emission of the next response in the chain on a discriminative basis. An extension seems to be warranted by the findings of the present experiment. This extension would summarize the effects of extinction of the preceding sequence $R_2 \rightarrow S^r$, insofar as it may be distinguished operationally for test purposes, on the strength of the sequence $S^D \cdot R_1$, when the stimuli in each sequence are identical.

The relationship between identical response members of a reflex chain may be formulated in the following terms: when both are governed by the same stimulus ($S^{r,D}$), their rates and totals will be equally increased by the continued availability of the stimulus, but when either is produced by the stimulus without completion of the chain and ultimate reinforcement, the amount of additional responding remaining under the control of the stimulus becomes smaller and smaller. When the responses differ in some respect, a similar interdependence may be predicted, although the rate and number of each may differ as a function of several variables associated with the difference in definition or specification.

It should be possible to obtain a further substantiation and generalization of the latter formulation by the study of an intact chain which includes at least two differing but recorded responses. Let the initial member of the chain produce the S^D for the next response; in the presence of this S^D, let the next response be reinforced periodically, at a fixed interval. The rate, then, or the extinction ratio (number of unreinforced responses emitted per reinforcement) serves as a measure of the strength of this later response. The rate of the response under these conditions can then be compared with the rate of the same response when its S^D is presented arbitrarily by the experimenter rather than as a consequence of an initial member of a chain.

Filter versus reserve.—When a response is increased in its rate of emission by being allowed to produce a stimulus which has previously set the occasion for the reinforcement of another response, Skinner speaks of 'secondary reinforcement.' This implies an increment in his hypothetical variable, the 'reserve.' Precisely defined, this reserve consists of the total number of unreinforced responses which will thenceforth be emitted; since a test period cannot be extended to an infinite duration, this is not measurable. However, Skinner further suggests that this reserve is diminished by each successive unreinforced response and that the rate of emission is a function of the size of the remaining reserve. Thus, the rate approaches the unconditioned or operant level of emission as an asymptote; when this asymptote is very nearly reached, the total number of responses which have been emitted serves as a fairly accurate measure of the reserve, although the criterion itself is not completely determinate. Such a measure has frequently been termed 'resistance to extinction.'

In the case of the discriminative use of a stimulus during extinction,

Skinner is inclined to believe that it acts as a 'filter.' That is, it affects the initial rate of emission independently of the reserve. If this is granted as an assumption, the rate is no longer a lawful function of the remaining reserve, and the approach of this rate to an asymptote is no longer so satisfactory a criterion for the duration of a test period designed to measure the reserve. Once more the reserve seems to be removed from the realm of exact measurement.

There remains, however, the possibility of using resistance to extinction as an important dependent variable in its own right. The question then becomes whether the presentation of the stimulus, either discriminatively or as a secondary reinforcer, merely influences the initial rate of responding (acts as a 'filter') or whether it adds to the number of responses occurring prior to the approach of the rate to the operant level (resistance to extinction).

It may first be pointed out that the extinction curves produced in the present study by the discriminative or reinforcing use of the stimulus are the same. It would be contradictory, therefore, to maintain separate formulations of the two stimulus effects. A choice must be made.

The operant or unconditioned rates of responding of the individual rats used in the present experiment have not been obtained. However, it has generally been assumed that when there is no known distortion of the results due to such factors as changes in drive or in emotional state resistance to extinction can be adequately predicted from the total number of responses emitted during an extended period; most experimenters have accepted a duration considerably shorter than that of the present study. Since the data demonstrate a substantial and progressive difference between the total responses accompanying the presentation of the effective stimulus either as S^D or as S^r and the total responses emitted in the absence of the stimulus, it may be concluded: resistance to extinction is increased, and to the same degree, by the use of either schedule of stimulus administration. Consequently, the filter analogy is rejected.

Theorectical distinction.—Within the theoretical framework developed by Skinner, it has been customary to treat the discriminative stimulus (S^D) and the secondary reinforcing stimulus (S^r) as two distinct and largely independent functional categories. Yet it has long been recognized that there are certain establishing operations, at least, which are common to both uses of the stimulus. Indeed, a survey of the experimental literature at the present time reveals no concrete evidence of a difference in the variables which govern the acquisition of these properties by a stimulus.

The present study further indicates that when the stimulus is continuously available the effectiveness of the two schedules of stimulus administration will be equal and will remain equal during the process of extinction; that is, they vary together in effectiveness as a function of at least one of the governing variables. It is conceivable that this correspondence is an accidental occurrence and that a change in other parameters such as size or quality of reinforcement, number of reinforcements, duration of cycle, physical characteristics of the stimulus, etc., would destroy it. It seems more probable, however, that the correspondence is a general one, which will be maintained under a variety of circumstances, provided that the temporal schedule of administration is not itself manipulated.

This conclusion permits a considerable simplification in the theoretical treatment of discriminative and reinforcing stimuli. It is no longer necessary to assume that the two uses of the stimulus are differentially affected by variations in the training procedure whereby they acquire control over behavior. That is, it is no longer necessary to provide two separate theoretical accounts of the manner in which a stimulus acquires the power to raise the rate of response. Moreover, the expenditure of the influence of the stimulus under either schedule brings about an equal reduction in the effectiveness of the other. The previous distinction between two types of stimulus is reduced to a distinction between two categories of temporal schedule for the administration of the stimulus: in the case known as secondary reinforcement, the presentation of the stimulus follows a specified response and is dependent upon the occurrence of the response; in the discriminative case, the stimulus is presented independently by the experimenter, without any necessary temporal relationship to a given response.

It is to be expected, of course, that differing schedules of stimulus administration will sometimes have different effects on the amount of responding. For example, the stimulus might be presented at fixed intervals of time, rather than continuously. The duration of presentation might be varied. The production of the stimulus might be made dependent upon a fixed number of responses. Just as variations in the temporal schedule for the administration of a food pellet affect the rate of responding and resistance to extinction, so may variations in the administration of this conditioned type of stimulus produce quantitative variations in behavior. A distinction may here be drawn between procedures in which the production of the stimulus is dependent in some way on the emission of responses (S^r) and procedures in which the onset of the stimulus is in no way affected by the animal's behavior but is independently determined by the experimenter (S^D). It should be realized, however, that such a distinction is by no means exhaustive and that it is not always relevant to the number of responses which will be emitted. In the present study, for example, the rate and total are not affected. Thus, the distinction seems to be no more important in quantitative terms than many another distinction in administrative schedule.

Another aspect of the administration of the stimulus is the selective effect which it may exercise over the various responses in the organism's repertoire. When the presentation of the stimulus is made dependent upon instances of a certain class of responses, e.g., bar-pressing, it may be said to be selectively correlated with this response. Occasions when it closely follows or precedes instances of some other response category will be coincidental. On the other hand, when the presentation of the stimulus is not dependent upon the occurrence of a specified response, no temporal correlation is imposed and no independent selection is exercised by the experimenter. Its temporal contiguity to any response which may be emitted is determined by the animal's own behavior.

A problem which demands further investigation in this case is the prediction of which response will be strengthened, or, more precisely, what factors determine the relative strengthening of a variety of response categories. The occurrence of relatively fixed sequences of responding in chaining situations makes it evident that the response which has actually been reinforced in the presence of a particular S^D is the one which is most affected by this stimulus, but experiments by Estes (3, 4) and Walker (12) demonstrate that this is not the only response which may thus be strengthened. Among the possible factors which may determine the relative effects of an S^D on a variety of response are: (1) the degree of induction between each response and the specific response used in training; (2) the number and quality of previous reinforcements directly produced by each response; (3) the operant rate of each response; or (4) a combination of these variables.

Summary and Conclusions

Thirty-six rats were given 200 cycles of discrimination training on the bar-pressing response, half with light as the S^D, half with darkness. Following this training, the bar-pressing response was extinguished under three schedules for the administration of the S^D: (1) 16 rats were extinguished in the presence of the S^D, but each response was followed by three sec. of S^Δ; (2) 16 rats were extinguished in S^Δ, but each response was followed by three sec. of S^D; (3) four rats were extinguished without the presentation of S^D at any time. These procedures were considered tests of (1) the discriminative role of the stimulus, (2) the reinforcing role of the stimulus, and (3) rate of responding in the absence of the stimulus (control). The cumulative curves of the first two groups could not be reliably distinguished, although they differed substantially from that of the control group.

Half of each of the experimental groups was transferred after the third daily session to the alternative test procedure and extinction was continued for three more days for all animals. Again, the rats which had undergone a change of procedure responded at the same rate as those rats which had remained on the same procedure.

These results indicated that the efficacies of the discriminative and reinforcing roles of a stimulus covary under the establishing and abolishing operations which were utilized. Three further generalizations are indicated by these findings:

1. Skinner's Law of the Extinction of Chained Reflexes should be amplified to include the observation that the extinction of a sequence $R \rightarrow S^r$ may reduce the strength of the following sequence $S^D \cdot R$.

2. The suggestion that an S^D acts only as a 'filter' during extinction, governing the initial rate of emission but not the total number of responses emitted, conflicts with the implication that an S^r can be used to create a 'reserve' of responses. The data suggest that both usages affect total responding (resistance to extinction).

3. It no longer seems to be necessary to provide separate theoretical accounts of the manner in which a discriminative or reinforcing stimulus gains or loses the power to raise the rate of response. The previous distinction between the two types of stimulus appears to be reduced to a distinction between two categories of temporal schedule for the administration of the stimulus. This distinction is relevant chiefly to the determination of the distribution of the stimulus effects among the variety of responses in the organism's repertoire.

(Manuscript received August 4, 1949)

References

1. Bugelski, R. Extinction with and without sub-goal reinforcement. *J. comp. Psychol.*, 1938, **26**, 121–133.
2. Ellson, D. G. Quantitative studies of the interaction of simple habits. I. Recovery from specific and generalized effects of extinction. *J. exp. Psychol.*, 1938, **23**, 339–358.
3. Estes, W. K. Discriminative conditioning. I. A discriminative property of conditioned anticipation. *J. exp. Psychol.*, 1943, **32**, 150–155.
4. Estes, W. K. Discriminative conditioning. II. Effects of a Pavlovian conditioned stimulus upon a subsequently established operant response. *J. exp. Psychol.*, 1948, **38**, 173–177.
5. Fisher, R. A. *Statistical methods for research workers.* London: Oliver & Boyd, 1936.
6. Frick, F. C. An analysis of an operant discrimination. *J. Psychol.*, 1948, **26**, 93–123.
7. Frick, F. C., Schoenfeld, W. N., & Keller, F. S. Apparatus designed for intro-

ductory psychology at Columbia College. *Amer. J. Psychol.*, 1948, **61**, 409–414.

8. McNemar, Q. L. *Psychological statistics.* New York: Wiley, 1949.
9. Notterman, J. M. A study of some relations among aperiodic reinforcement, discrimination training and secondary reinforcement. (To be published)
10. Schoenfeld, W. N., Antonitis, J. J., & Bersh, P. J. A preliminary study of training conditions necessary for secondary reinforcement. *J. exp. Psychol.*, 1950, **40**, 40–45.
11. Skinner, B. F. *The behavior of organisms.* New York: Appleton-Century, 1938.
12. Walker, K. C. Effects of a discriminative stimulus transferred to a previously unassociated response. *J. exp. Psychol.*, 1942, **31**, 312–321.
13. Youtz, R. E. P. The weakening of one Thorndikian response following the extinction of another. *J. exp. Psychol.*, 1939, **24**, 294–304.

RESISTANCE TO EXTINCTION OF A CONDITIONED OPERANT AS RELATED TO DRIVE LEVEL AT REINFORCEMENT *

BY RAYMOND CORNELIUS STRASSBURGER

From the Psychology Laboratory, Columbia University

Introduction

The drive variable in relation to learning has been studied in a number of researches stemming from diverse theoretical formulations and employing varied experimental arrangements. Among these investigations several have been directed toward the dimension of intensity or level of drive as a factor in the learning process. The learning tasks employed in these drive studies have involved behavioral sequences ranging from complex patterns to simple response segments. Typical of a series of earlier investigations of maze learning in relation to strength of drive are the studies of Ligon (10) and of Tolman, Honzik, and Robinson (17) which demonstrate that facility in acquisition of the maze habit by rats tends to vary directly with the degree of hunger. Warner (19), using the obstruction method in which a conflict of drives is evoked, reports that the number of grid crossings by hungry rats rises to a peak somewhere between the second and fourth days of deprivation, and then declines.

While these earlier studies point to a probable relationship between drive and behavior, the mechanical arrangements and the experimental designs employed do not permit of a clear-cut segregation of the drive factor as a performance variable from this same factor as operating in the acquisition of the habit. The distinction between the utilization of what has already been learned and the acquisition of new modes of response has been emphasized by some theorists (8, 9, 18). It is a reasonable hypothesis that the effect of motivation is not uniform but varies with the distinct phase of the behavioral process under investigation. Consequently, experimental situations should be structured to permit the testing of the drive relationship in each behavioral aspect independently of the contaminating influence of other relationships which may exist.

Analysis of the behavioral process into the conditioning phase and the extinction phase provides a basis for such an experimental approach. Given a relatively simple segment of operant behavior, it is possible to set up a procedure which systematically varies the drive at the time of conditioning of the emitted response. By subsequently extinguishing the conditioned reflex under the same drive conditions, a measure of the strength of conditioning, and of the effect of drive upon the conditioning, is obtained. Thus the quantitative relation existing between drive and reinforcement may be determined.

It was in these terms that Finan (2) formulated an experimental study of the effects of reinforcement under different degrees of food deprivation upon the strength of lever-pressing responses in rats. In this investigation groups of animals which had been given 30 reinforcements under drive conditions produced by varying intervals of fasting were later subjected to experimental extinction of the conditioned pressing response under uniform conditions of hunger. Using resistance to extinction as the measure of the original excitatory tendency produced by reinforcement, Finan reports evidence of a relationship between strength of conditioning and

* Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Faculty of Pure Science, Columbia University. The writer acknowledges his indebtedness to Professors W. N. Schoenfeld, R. F. Hefferline, and C. J. Warden for invaluable assistance in the formulation of this research, and for constant encouragement and guidance in carrying it to completion.

intensity of hunger prevailing at the time of reinforcement. Habit strength increases from the relatively low level produced by conditioning after one hour of deprivation to a maximum at 12 hours of deprivation, and then declines at a negatively accelerated rate through the 24- and 48-hour deprivation points.

Among the several deprivation groups the differences in strength of conditioning (as measured by the number of extinction responses) between one hour and the other periods are significant; the differences among the other periods are not significant. Data based on extinction time and rate generally show less reliable differences.

The conception of reflex strength as a positive function of drive intensity at the time of conditioning is not confirmed in a subsequent research by Kendler (6), who also employed the same operant of lever-pressing in a Skinner apparatus. In this study of drive interaction the responses were reinforced in two groups of rats under 22 hours of food deprivation but with different degrees of thirst, i.e., after 12- and 22-hour periods of water deprivation respectively. Extinction of the response under uniform drive conditions produced no evidence of a difference in reflex strength attributable to the initial variation in the thirst drive. Kendler concludes in view of these results that the intensity of drive at the time of acquisition is not a factor in habit growth, a conclusion which he finds consistent with the theoretical formulation of Hull (4). These contradictory findings point up the need for further data.[1]

In broad outline, the research here reported is essentially a replication of the work of Finan, and adheres closely to the experimental design of this earlier investigation. Some minor deviations incorporated in the present procedure have been dictated by the exigencies of the experimental situation or by the inability to duplicate therein the exact conditions attained by Finan. In a major respect, however, it represents an extension and development of his research. Whereas Finan restricted his attempt to determine a four-point curve describing the relationship between drive and strength of conditioning to an experimental situation involving 30 reinforcements, the present study would assess the character of the function for different degrees of reflex strength as induced by varying numbers of reinforcements. It is manifestly desirable to make the experimental test over a broad range of values of the experimental variable, in order that any inference derived therefrom may not be hampered by overly restricted quantitative data.

The drive operations selected as the basic experimental variables in this study are those of feeding and fasting. They are readily variable along a temporal continuum, and provide a drive factor which may be related to changes in reflex strength under conditions of reinforcement. An indirect measure of habit strength is available in terms of resistance to extinction of the response, that is, when it is continuously emitted without reinforcement. The experimental problems here proposed may then be stated as follows: How does resistance to extinction of a conditioned operant vary with the length of the interval of fasting preceding original conditioning of that response? Further, does the nature of the relationship between time of deprivation and response strength vary with the number of reinforcements given in the conditioning process?

An experimental design appropriate to the testing of these hypotheses requires conditioning or training under varying periods of fasting, among several different reinforcement groups,

[1] A recent paper by Bradley Reynolds bearing on this problem (cf. this JOURNAL, 1949, 39, 296–305) has appeared too late for review here.

followed by subsequent extinction sessions carried out in all cases under uniform conditions of hunger. To assess the effect of drive operations on as broad a base as feasible a wide range of deprivation periods is desirable. In this study the influence of extremely short intervals of fasting (½ and one hour) as well as periods of relatively long duration (47 hours) are investigated at levels of habit strength which extend from that produced by 30 reinforcements to that resulting from the conditioning effect of a single reinforcement.

Apparatus and Subjects

The basic apparatus of the type designed by Skinner (16) for studying the conditioned bar-pressing response in white rats, as modified and developed in the Columbia Laboratory (3), was used in this study. The chief feature of this equipment is a combination living cage and experimental box into which the response lever may be introduced during the experimental session, so that the animal is housed and maintained in the situation in which experimentation takes place. Emotional disturbances resulting from transfer of the animal to the novel surroundings of a strange experimental apparatus are consequently avoided. Permanent records of the animal's behavior in the form of number and temporal distribution of the responses were made available by means of a kymograph rotating at the rate of one revolution in 6 hours. On its horizontal drum a writing point traces a step-like cumulative record of the lever pressings. A waxed paper polygraph with a signal marker activated by the bar supplied an additional record of the time of occurrence of each response. The complete apparatus is fully described in the above citation (3).

The experimental chamber was a quiet, darkened room from which external disturbances were excluded. A small electric bulb of three watts suspended about two in. above the experimental box illuminated its interior and provided a minimal yet adequate source of light for the experimenter to observe the animal and to perform the necessary experimental operations.

Male Wistar rats of the age range 90–110 days were employed in this research. A total of 170 animals entered tentatively into the investigation, although usable data for analysis were obtained from a lesser number. Of the original population, 6 rats were eliminated through death or illness during the course of an experimental unit, 19 were discarded because of bar-biting, 7 failed to eat the requisite number of pellets provided as reinforcements, and 4 did not attain the specified criterion of extinction in the time allowed—one hour. The relatively large number of 19 animals discarded because of biting or gnawing on the response lever during conditioning or extinction was not concentrated in any one deprivation group, but was distributed among the several drive conditions. These eliminations reduced the number of animals from which satisfactory data were available to 134, comprising samples of 10–12 rats in each of the deprivation sub-groups within the three different reinforcement conditions, and a small independent group of five animals used as a control to determine the rate of unconditioned bar-pressing.

Procedure

Each experimental unit, consisting of the three main phases of preliminary feeding and training, conditioning, and extinction, extended over a period of nine days (11 days in the case of the single reinforcement group where a modification in procedure was introduced). During this time each animal was housed in its own living-working cage and maintained under the usual animal colony conditions. The diet consisted of a standard, prepared animal food, Purina Laboratory Chow, in rough-ground, dry form, with a constantly available supply of water.

Days 1–4.—The experimental cycle was initiated by a program of restricted once-a-day feeding to establish a hunger rhythm. Over a period of four days the standard ration of 10 gm. was made available to the animal daily for one hour, at the end of which any remaining food was removed. The rats grew well and gained weight on this ration. All subsequent training and experimental sessions occurred at the regular feeding time for each animal, coincidental with the peak of the hunger cycle.

Day 5.—This phase of preliminary training was designed to habituate the animals to the serial presentation of food pellets under actual experimental conditions. These pellets, which were employed as reinforcing stimuli in the later conditioning session, were prepared from the standard ration according to a procedure fully described elsewhere (15). The hard, cylindrical masses were approximately 0.7 cm. long by 0.3 cm. in diameter, and weighed about 0.07 gm. The training procedure was carried out in the darkened experimental room under the previously described conditions which were standard for conditioning and extinction sessions as well as for the training series. Coincidental with the usual feeding time for the animal which had

fasted for the previous 23 hours, a pellet was dropped through the chute leading to the food tray inside of the experimental box. Accompanying the delivery of the pellet was an auditory cue provided by the recording apparatus. This discriminative stimulus was identical with that subsequently produced in the conditioning trials by the animal's responding. After the first pellet was consumed by the animal, the routine was repeated until 40 pellets were presented and consumed in this fashion. This period of training proved adequate to insure the animal's reaction to the discriminative auditory stimulus and prompt conditioning when the bar was introduced in the subsequent conditioning trials carried out two days later. The effectiveness of this preliminary training routine is evident in the rapidity with which a maximum rate of responding to the lever was attained in conditioning. Reaction to the bar quickly followed its introduction into the cage in most cases, and the first lever-pressing and its reinforcement were consummated in a period of time averaging about one min.

After the termination of the training session the balance of the daily food ration of 10 gm. was made available to the animal in the usual way. The training and feeding routines occupied about one hour.

Day 6.—The animals were fed at the customary time, with the exception of those in the 47-hour and in the 11-hour deprivation groups. In the case of the former the usual feeding was omitted and their deprivation continued through the 6th day, so that a period of 47 hours of fasting might precede conditioning on the 7th day. Animals to be conditioned after 11 hours of fasting received their 6th day feeding 12 hours beyond the usual time, thus providing a deprivation period of 11 hours prior to conditioning on the next day. At the last feeding preceding conditioning the animals were offered food beyond the standard ration to insure satiation in all groups.

Day 7.—This was the day on which conditioning occurred. The animals, according to their deprivation group, were fed to satiation ½, 1, and 4 hours prior to their usual feeding time. The conditioning procedure in all cases was begun at the time at which the customary feeding and the previous training had taken place.

The conditioning procedure involved transferring the rat in his work box to the experimental room where a period of five min. for adaptation was allowed before experimentation began. Then the response lever was introduced about one cm. through the slot at the end of the box directly over the food tray. Subsequent investigatory responses of the animal would occasion contact with the bar, which was responsive to a downward pressure of 10 gm. The movement of the lever actuated the recording apparatus, producing the auditory signal to which the preliminary training had already made the animal responsive. Simultaneously a pellet was released by *E* into the food tray to reinforce the response. Conditioning was continued in this way until the specified number of reinforced bar-pressings was obtained. At the close of the conditioning trials, food was made available to the animal for a period of one hour.

Table I specifies the various hunger periods at which conditioning to the level represented by the different numbers of reinforcements was accomplished.

TABLE I

DEPRIVATION AND REINFORCEMENT CONDITIONS IN THE SEVERAL EXPERIMENTAL GROUPS

Experimental Group	Number of Reinforcements	Hours of Deprivation					
		½	1	4	11	23	47
A	30	—	—	11*	12	10	11
B	10	11	10	11	—	12	—
C	1	—	10	—	10	11	10

* The figures represent the number of animals in each of the specified deprivation sub-groups.

Although the present study follows closely the procedure formulated by Finan, several extensions of his experimental plan are evident. Among the deprivation intervals employed by Finan was that of one hour. Experience clearly demonstrated the difficulty of securing 30 reinforced responses from animals conditioned after this short period of fasting. Since a reinforced response required the consumption of a food pellet directly upon depressing of the lever, animals fasting for a period of only one hour were not sufficiently hungry to maintain the routine for 30 trials. As a result a somewhat longer interval of four hours was selected for the 'low-drive' group at this level of reinforcement. The remaining deprivation intervals are designated 11, 23, and 47, rather than 12, 24, and 48 hours as a consequence of reckoning the period of fasting from the end rather than the beginning of the feeding session which lasted one hour.

Hunger intervals in Group B were selected to explore the effect of brief deprivation periods, approaching zero hunger drive, upon the strength of conditioning. A relatively small number of reinforcements was necessary to make this possible, and a total of 10 was chosen. Some overlapping of the fasting periods with those

prevailing in Group A, the 30 reinforcement group, was judged desirable in order to make limited comparison between the groups.

Skinner (16) has reported conditioning, under optimal circumstances, with a single reinforcement, where a complete and maximal change in the rate of responding is produced. It has, moreover, been asserted that the role of motivation has not been explored sufficiently at the level of 'bare' acquisition (8). These considerations dictated the choice of an experimental condition for Group C in which only a single bar-pressing was reinforced before the extinction process was begun.

Day 8.—Animals in all of the deprivation groups, conditioned 24 hours earlier, were fed at the usual time. The purpose was to restore the hunger rhythm previously interrupted, and to insure the equating of the hunger drive for the extinction session of the following day.

Day 9.—The extinction phase of the experiment was carried out with uniform conditions of hunger prevailing in all groups. At the time of extinction 23 hours had elapsed since the termination of the previous feeding session. During this extinction process all bar-pressing responses were unreinforced. Extinction was continued until a 10-minute interval without response was reached. If this criterion was not attained within a period of one hour the process was terminated at that time. This procedure, which was observed with Group A, was modified in the case of the other experimental groups, in which all animals were subjected to one hour of extinction, when the process was assumed to be accomplished. This alteration of the experimental plan was considered advisable in order to provide additional data for possible analysis in the total number of responses emitted during the one-hour extinction interval. Within this period, data for various criteria of 2, 5, 8, and 10 min. of no response also were tabulated and analyzed.

This general experimental plan was altered somewhat with Group C, the single reinforcement group. After the usual four days of preliminary feeding, on the 5th day the bar was made available to the unconditioned, hungry rat, and a record of 'unconditioned' pressings for a period of 1 hour was secured. These responses were neither followed by the delivery of a food pellet nor by the auditory signal. The data thus obtained provide a basis for assessing the effect of the single reinforcement in the subsequent conditioning. This would be evidenced in an increase in the number of responses during extinction over that emitted in an equal period of time prior to reinforcement. In addition, such data serve to demonstrate the comparability of the deprivation groups in respect to the strength of the operant before conditioning. Keller and Schoenfeld (5) have stressed the desirability of securing so-called 'operant level' readings in experiments involving the comparison of groups subjected to varying experimental conditions, in view of the demonstrated variability in the rate of unconditioned bar-pressing among animals (cf. also 14). Such precaution is particularly important in the case of small samples, and desirable even with large groups. The random assignment of animals to the different deprivation groups followed in this experiment, however, does not appear unjustified in view of the remarkably close correspondence among the group means obtained without selection in the case of Group C. Here the mean number of bar-pressings for an hour period, without prior reinforcement, in the four experimental groups showed little variations: 35.6, 36.2, 34.5, and 35.7. The standard deviations were somewhat more variable: 13.0, 9.86, 15.4, and 10.9, respectively. In another independent group of five animals employed as a control, and again randomly selected, unconditioned bar-pressing occurred at the mean rate of 30.6 responses per hour, with a standard deivation of 7.7. The median in this group was 34.0 responses.

Apart from this modification in the experimental sequence to obtain data on unconditioned responding to the lever, the animals in the single reinforcement group were subjected to the same routine in respect to conditioning and extinction as were the animals in the other reinforcement groups designated as A and B.

Results

Conditioning

Since it is conceivable that the effect of variation in the length of the deprivation interval preceding conditioning may be manifest in some aspect of the conditioning process, an analysis of the data derived therefrom was undertaken. From the recorder tapes which register the temporal distribution of responses in relation to the point of introduction of the lever, the interval of time from the initial availability of the bar to any specified response in the reinforcement series can be determined. Finan has designated the total time required to complete the reinforcement process 'learning time,' and he regards it as an

index of the facility with which conditioning occurs. The somewhat ambiguous nature of this measure, however, must be recognized. Broken down into its component elements, it comprises (a) the time to establish initial contact with the bar followed by the first reinforcement, and (b) the subsequent bar-pressings up to the stipulated number of reinforcements. Response latency in the sense of time to the first reinforcment, however, is too dependent upon incidental and chance factors, such as position of the animal in the box when the bar is introduced, among others, to be of much significance in assessing the effect of drive variation. Reinforced responses after the first rapidly attain a maximum rate, and the time then becomes largely a measure of 'eating rate.' Previous studies (16) have demonstrated that the slopes of eating curves with regular reinforcement are not clear functions of drive level, as is true in the case of periodic reinforcement. Although the somewhat ambiguous character of such a measure is recognized, and although its significance is not thoroughly clear, to permit comparison with Finan's results, data for 'learning time' are reported here. In Table II appear the mean times to complete the first 10 reinforcements and the total 30 reinforcements in Group A, and the total 10 reinforcements in Group B, for the several deprivation conditions. No data for the single reinforcement group, Group C, are presented since the time up to the first reinforcement can be of little significance for the reasons cited above.

There is no conclusive evidence in the data of Table II that the rate of reinforcement is a function of the degree of hunger existing at that time. Superficially it would appear that animals trained after fasting 11 hours and 23 hours tend to attain a designated number of reinforcements more rapidly than those conditioned after longer or shorter intervals of deprivation. This tendency is most clearly evidenced in Group A, where a period of 47 hours of fasting produces a manifestly reduced overall rate of response and ingestion for training to the level of 30 reinforcements. The differences in mean training time for the animals in this group and those in the 11-hour and the 23-hour groups reach a high level of statistical significance, as Table III indicates. For 10 reinforcements, the time required for training by the 23-hour animals is also significantly lower as compared with the one hour and possibly the four-hour groups. As a tentative conclusion it appears that a 23-hour

TABLE II

Effect of Hunger on Gross Training Time

Time in Min.	Deprivation Time in Hours					
	½	1	4	11	23	47
Group A						
Mean time, 1st 10 reinf.	—	—	5.0	4.1	4.5	4.9
S.D.	—	—	2.5	1.2	2.1	1.6
Mean time, 30 reinf.	—	—	11.7	9.3	9.3	12.7
S.D.	—	—	4.7	1.8	3.1	3.7
Group B						
Mean time, 10 reinf.	4.3	5.6	5.4	—	3.4	—
S.D.	1.5	2.7	3.2	—	1.2	—

TABLE III

RELIABILITY OF DIFFERENCES IN MEANS OF GROSS TRAINING TIME (P)

Mean Time	½ hr. vs. 1 hr.	½ hr. vs. 4 hr.	½ hr. vs. 23 hr.	1 hr. vs. 4 hr.	1 hr. vs. 23 hr.	4 hr. vs. 11 hr.	4 hr. vs. 23 hr.	4 hr. vs. 47 hr.	11 hr. vs. 23 hr.	11 hr. vs. 47 hr.	23 hr. vs. 47 hr.
Group A											
1st 10 reinf.	—	—	—	—	—	>.30	>.60	>.90	>.60	>.20	>.60
30 reinf.	—	—	—	—	—	>.10	>.20	>.05	1.00	.01-.02	.02-.05
Group B											
10 reinf.	>.20	>.30	>.10	>.80	.02-.05	—	>.05	—	—	—	—

period, identical with the established feeding rhythm, is productive of most rapid training, while the longer period of 47 hours is relatively inefficient.

With respect to so-called 'learning time,' Finan reported ". . . a small but consistent tendency for more highly motivated groups to consume less time in learning to depress the bar to the 30-response criterion . . ." (2, p. 126). The reliabilities of the differences in his data, however, are too low to warrant substantial confidence in this conclusion. Other investigators (1, 16) have discovered that the relation between rate of feeding and hunger is a complex one, and Skinner has stated that ". . . a rat does not always eat more rapidly the more depleted it becomes" (16, p. 352). Thus the implications of the present study respecting the effect of hunger upon the gross time in training are not inconsistent with the results of prior research.

Extinction

A. *Criteria of extinction.*—The use of resistance to extinction as a measure of reflex strength implies the availability and application of some criterion of extinction (11). Some investigators have favored a standard interval of unreinforced responding, generally one hour, at the end of which a near-zero rate is assumed to exist (16). Others have employed as a criterion a shorter period of no responding, generally an interval of 5 or 10 min. with the lever constantly available to the animal (2, 7, 20), or the absence of response during a sequence of there successive, relatively brief, insertions and removals of the response lever (6). In the present inquiry data are recorded respecting number, time and rate of responding up to intervals of 2, 5, 8, and 10 min. of no response, and with the exception of Group A, for total number of responses during the one hour extinction period.

Correlations among the several criteria were computed and are reported in Table IV. In determining these correlations the several deprivation groups under each reinforcement condition are combined since the evidence indicates that the number of extinction responses is invariant with different drive levels in conditioning. There is, moreover, some unavoidable selection in view of the failure of a number of animals to attain

TABLE IV

PRODUCT-MOMENT CORRELATION OF NUMBER OF RESPONSES UP TO DESIGNATED INTERVALS OF CESSATION OF RESPONSE WITH TOTAL NUMBER OF RESPONSES FOR ONE HOUR OF EXTINCTION IN THE DIFFERENT REINFORCEMENT GROUPS

Criterion	Total Number of Responses in the Different Reinforcement Groups		
	1 Reinf.	10 Reinf.	30 Reinf.
2 min. no resp.	.503 (41)	.951 (44)	.852* (33)
5	.815 (41)	.910 (44)	.929* (33)
8	.907 (40)	.896 (39)	.971* (33)
10	.917 (37)	.897 (31)	—

* In the case of the 30-reinforcement group correlations are between the designated criteria and the number of responses up to 10 minutes of no response. *N* is given in parentheses.

the more rigid ceriteria of 8 and 10 min., although reaching to two- and five-min. standards within the time allotted for extinction. This accounts for the variation in the number of cases contributing to the correlations within a given reinforcement condition. There is no drop in *N* within the 30-reinforcement group, since a total of 33 animals attained the 10-min. extinction criterion with which correlations are taken here.

Examination of the coefficients of correlation reported in Table IV reveals a high degree of correspondence among the different criteria. The selection of a five-min. criterion for the final analysis of results seems justified in view of these experimental findings. Such a choice has the merit of increasing the amount of data for interpretative analysis, and of being consistent with the practice of other investigators. Moreover, analyses of the data in terms of the other criteria (not reported here) yield results in close conformity to those derived when the five-min. standard is applied.

B. *Resistance to extinction.*—Quantitative measures of the degree of resistance to extinction of a response tendency are available in the number of responses emitted during extinction, in the time required for the process, and in the overall rate of responding through the extinction period. It is in these aspects of the response process that an indirect measure of reflex strength in relation to variation in drive at the time of reinforcement may be found. Previous investigators have tended to favor the number of unreinforced responses during extinction as providing the most reliable and sensitive measure of reinforcement, although both time and rate have been utilized (**2, 20, 21**). All three indices figure in the analysis of results of the present study. The effect of variation in hunger at reinforcement on resistance to extinction as expressed in terms of these measures is shown in Tables V through VII. Table VIII reports *t*-test reliabilities of the differences among the means for the several measures of reinforcement.

In Table V are assembled and summarized the results describing the effect of differential drive at training upon the number of responses in extinction for each of the three reinforcement groups. The increase in reflex strength resulting from a single reinforcement is remarkably constant in spite of a wide variation in the conditioning drive. The largest mean dif-

TABLE V

NUMBER OF RESPONSES TO EXTINCTION *

Reinforcement Group	Hours of Deprivation at Conditioning					
	½	1	4	11	23	47
30 reinforcements						
Mean no. responses	—	—	128.3	111.4	128.1	123.5
S.D.	—	—	57.6	43.3	37.9	56.8
N	—	—	11	12	10	11
10 reinforcements						
Mean no. responses	74.4	94.3	68.5	—	99.4	—
S.D.	27.2	52.0	26.4	—	57.1	—
N	11	10	11	—	12	—
1 reinforcement						
Mean no. responses	—	42.4	—	39.0	45.0	51.8
S.D.	—	22.0	—	20.1	29.9	30.0
N	—	10	—	10	11	10

* Criterion is 5-min. interval of no response.

TABLE VI

TIME TO EXTINCTION * IN MINUTES

Reinforcement Group	Hours of Deprivation at Conditioning					
	½	1	4	11	23	47
30 reinforcements						
Mean time in minutes	—	—	25.5	22.2	21.0	21.5
S.D.	—	—	14.1	8.1	4.5	8.3
N	—	—	11	12	10	11
10 reinforcements						
Mean time in minutes	12.9	19.4	17.8	—	21.3	—
S.D.	4.4	11.9	6.3	—	14.0	—
N	11	10	11	—	12	—
1 reinforcement						
Mean time in minutes	—	15.4	—	14.1	14.0	17.9
S.D.	—	6.8	—	8.8	7.1	10.2
N	—	10	—	10	11	10

* Criterion is 5-min. interval of no response.

ference in the number of responses is 12.8 which exists between the 47-hour and 11-hour groups. There is marked variability, however, within each drive group, and the *P*-value here is >.30. More substantial absolute differences appear where the number of reinforcements is greater. With 10 reinforcements in training, reflex strength at extinction is lowest in the group conditioned after four hours of fasting, and highest in the 23-hour group. Intra-group dispersion of scores is large, and the mean difference of 30.9 has a *P*-value >.10. Where 30 reinforcements are given, lowest resistance to extinction appears in the 11-hour animals, while conditioning after shorter and after longer intervals of deprivation produces greater reflex strength. Wide dispersion is again apparent. The largest difference reveals a *P* of >.30. There is here obviously no conclusive, unequivocal

TABLE VII

RATE OF RESPONSE TO EXTINCTION *

Reinforcement Group	Hours of Deprivation at Conditioning					
	½	1	4	11	23	47
30 reinforcements						
Mean rate of response	—	—	5.9	5.2	6.1	6.1
S.D.	—	—	2.3	1.6	0.9	2.4
N	—	—	11	12	10	11
10 reinforcements						
Mean rate of response	5.8	5.5	4.0	—	5.3	—
S.D.	1.6	2.5	1.4	—	2.3	—
N	11	10	11	—	12	—
1 reinforcement						
Mean rate of response	—	2.8	—	3.6	3.4	2.9
S.D.	—	0.9	—	2.0	1.3	0.7
N	—	10	—	10	11	10

* Rate is computed by dividing the number of responses in extinction by the time lapsed. Criterion is 5-min. interval of no response.

TABLE VIII

RELIABILITY OF DIFFERENCES AMONG MEASURES OF RESISTANCE TO EXTINCTION (*P*-values from *t*-tests)

Privation Groups	Number of Responses	Extinction Time	Rate of Response
30 reinforcements			
4 hr. vs. 11 hr.	>.40	>.50	>.40
4 hr. vs. 23 hr.	>.90	>.20	>.80
4 hr. vs. 47 hr.	>.90	>.40	>.80
11 hr. vs. 23 hr.	>.30	>.60	>.10
11 hr. vs. 47 hr.	>.50	>.80	>.30
23 hr. vs. 47 hr.	>.80	>.80	1.00
10 reinforcements			
½ hr. vs. 1 hr.	>.30	>.10	>.70
½ hr. vs. 4 hr.	>.60	>.05	.01–.02
½ hr. vs. 23 hr.	>.20	>.05	>.50
1 hr. vs. 4 hr.	>.20	>.70	>.10
1 hr. vs. 23 hr.	>.80	>.70	>.80
4 hr. vs. 23 hr.	>.10	>.40	>.10
1 reinforcement			
1 hr. vs. 11 hr.	>.70	>.70	>.20
1 hr. vs. 23 hr.	>.80	>.60	>.20
1 hr. vs. 47 hr.	>.40	>.50	>.70
11 hr. vs. 23 hr.	>.60	>.90	>.70
11 hr. vs. 47 hr.	>.30	>.40	>.30
23 hr. vs. 47 hr.	>.60	>.30	>.20

demonstration of the influence of differential drive at any level of reinforcement. There is no basis for assuming that the mean number of responses in the different groups represent other than random samples from the same population, and the conclusion of Finan pointing to such a relationship is not supported.

The effect of differential training as regards the number of reinforcements is clearly evident in the results, confirming the findings reported earlier by Williams (**20**). Disregarding the conditions of deprivation, the mean number of responses in extinction up to the 5-minute criterion increases from 44.6 in the single reinforcement group to 84.3 where 10 reinforcement are given in training, and to 122.5 where conditioning involves 30 reinforcements.

The mean time required to attain the criterion of extinction shows little variation in the different deprivation groups within any given reinforcement condition. From Table VI there is no clear evidence of a trend toward greater resistance to extinction of the reflexes reinforced under more intense drive. At 30 reinforcements reflex strength is greatest in the animals reinforced after only four hours of fasting, although with fewer reinforcements, i.e., 1 and 10, the 'higher drive' groups show slightly greater resistance to extinction. The mean differences are not significant. This is true even in the case of the ½-hour sub-group at the level of 10 reinforcements which shows substantially shorter extinction time to the five-min. criterion than the other drive sub-groups. If the 10-min. criterion is applied, this difference is not maintained. On the whole, in terms of extinction time, there is little evidence of a trend toward a consistent relation between strength of conditioning and hunger.

The measure called 'rate' as used here is a simple ratio of the number of responses in extinction to the time lapsed. This constitutes what might be called the 'overall rate' of response and is a crude measure. The data for rate of response over the extinction period in the different deprivation groups appear in Table VIII. The trends here necessarily parallel in some respects those evident in the case where number of responses is the criterion, because of the manner in which the data are computed. With 30 reinforcements in training, rate of response is lowest in the 11-hour deprivation group, and higher where conditioning occurs after shorter and after longer intervals of fasting. Where 10 reinforcements are given, the four-hour period of deprivation produces the lowest resistance to extinction in terms of rate. For a single

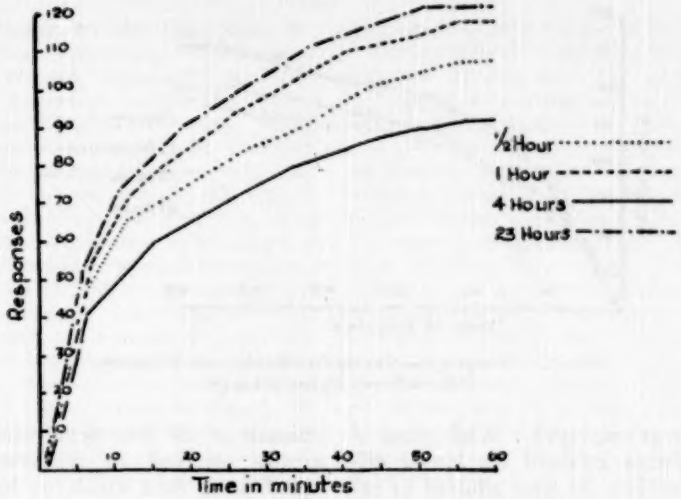

FIG. 1. Average extinction curves following ten reinforcements under different degrees of hunger.

reinforcement, however, the 11-hour interval appears to be more effective than either a shorter or a longer period of fasting. These differences are not reliable, although in the 10-reinforcement group the difference between ½ hour and four hours of fasting shows a *P* between .01 and .02. Rate of response in extinction, computed this way is, therefore, not apparently a function of the degree of hunger prevailing at the time of conditioning.

Cumulating the number of responses by five-min. intervals and averaging these data in each of the deprivation groups provides average extinction curves which illustrate the rate of response over the extinction period. Since a full hour was allowed the process in the cases of the one- and 10-reinforcement groups, such curves are available for these groups. However, with 30 reinforcements, extinction to the 10-min. criterion in a substantial number of cases was terminated before an hour had lapsed. Figs. 1 and 2 present average extinction curves following 10 reinforcements and one reinforcement, respectively, under different degrees of hunger. While inspection of the curves for a constant number of reinforcements under different drive levels reveals apparently different asymptotes, in view of the absence of any statistical confirmation of the reliability of the differences among the total number of responses for one hour, it may be assumed that in reality the asymptotes are the same. As regards the rates of approach to these asymptotes, there appears to be no reason for considering them other than invariant under the several drive conditions.

C. *Failure to Reach Criterion.*—The original plan of the present experiment contemplated carrying out the extinction process until an interval of 10 min. of no responding had been reached, assuming that occurred within one hour. For reasons already cited, however, in the final analysis of the data, a five-min. cri-

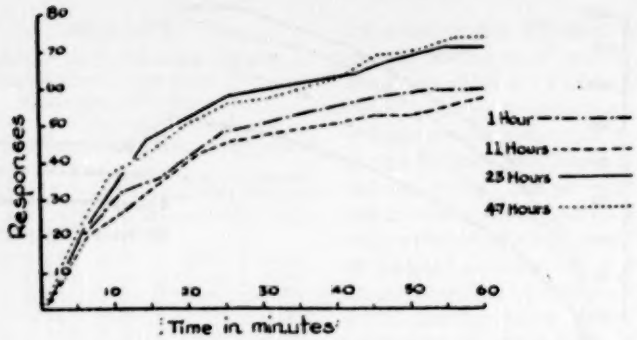

FIG. 2. Average extinction curves following one reinforcement under different degrees of hunger.

terion was employed. While most of the animals attained the longer criterion within the time allotted to extinction, there were some which reached the five-min. interval during which no responses were made, yet failed to attain the 10-min. criterion. The numbers of animals in the various deprivation groups which did not meet the longer criterion of extinction yield no evidence of any consistent trend reflecting an increase in resistance to extinction of the response with a lengthening of the fasting period prior to conditioning. Failure to extinguish in terms of a longer interval of no responding is probably a function of the number of reinforcements, and is apparently unrelated to the drive strength at conditioning within any reinforcement group.

DISCUSSION

In contrast with Finan's conclusions, the present study demonstrates that the effect of reinforcement upon the strength of the bar-pressing response is invariant within the range of fasting intervals covering ½ hour to 47 hours. The most significant comparison between the results of the two studies would be in the case of Group A of the present experiment where the animals of the four deprivation subgroups received 30 reinforcements, and Finan's data which are based on the same number of reinforcements. Animals fasting for only one hour prior to the conditioning session would not eat 30 pellets in training, however, and as a result the longer interval of four hours was substituted for the one-hour period which Finan used. It is noteworthy that he reports no difficulty in getting his animals to consume 30 pellets under these deprivation conditions. The modification here introduced, however, rules out a comparison with Finan's conclusion that the difference between the one-hour hunger period and the other periods of deprivation were reliable. Yet in Group B of the present experiment, where the number of reinforcements was reduced to 10 to permit the testing of fasting intervals of one hour and less, and thus to allow relating the results to Finan's study, the absence of a relation between degree of hunger and resistance to extinction is confirmed. The reinforcing effect of a single pellet, furthermore, is demonstrated to be invariable in the case of the animals in Group C, whether the deprivation interval is one hour or 47 hours.

The reasons for the discrepancy between these results and Finan's conclusions are not immediately apparent. Several differences in the two experimental conditions may be noted. While Finan used automatic feeding in which the delivery of the pellets was controlled exclusively by the animal, this experiment involved manual feeding so as to insure the individual reinforcement of each bar-pressing response through the consequent and immediate consumption of the food pellet. No comment is made in his report but it is assumed that each pellet was eaten by the animal following each one of the 30 acquisition responses. Although Finan presents no data on unconditioned bar-pressing in his animals, the groups are sufficiently large (approximately 30 animals) to justify the assumption of comparability in this respect. However, operant level data would be desirable in an experiment of this type. Again, although the procedure here did not involve the transfer of the animal to the experimental box at each session as was true of Finan's routine, it is unlikely that this difference was a major contributing factor to the results. Thus no ready explanation of the variation in the experimental findings presents itself.

The earlier studies briefly cited in the introduction point to a variation in maze performance and obstruction-box behavior with altered drive. Yet such evidence does not clearly demonstrate what it is in the total performance that varies with drive. Is it the strength of the habit itself, or the degree of utilization of the habit, or the extent of general activity incidental to the establishment or strengthening of the habit, or some other factor? The experimental structures of these studies do not permit a clear-cut segregation of these several aspects of the learning process. A distinction between the reinforcement phase and the extinction phase of the learning process needs to be made, and the several phases independently studied in relation to drive. The apparatus used in the present study makes possible this separation and independent evaluation.

There is no reason for regarding bar-pressing as a unique learning situation. It differs chiefly from the maze in that it does not demand from the animal as large an amount of response chaining or perhaps spatial exploration. There are no a priori grounds, therefore, for assuming that the quantitative laws describing behavior change in bar-pressing should be any different from those applicable to the more traditional learning situations. It is possible, for example, that the difference as regards the role of drive in maze learning and lever-pressing conditioning arises because the maze demands increased and sustained activity to produce the response elements which constitute the behavior pattern that is reinforced. In maze learning the effect of reinforcement as manifested in increased habit strength is not well differentiated from behavior potentiality or performance. Hence, the evidence from such studies may serve merely to demonstrate drive as a parameter of habit utilization, or response extinction. Such a conclusion is not inconsistent with the present findings.

The implications of the present study as regards the effect of drive variation upon the response rate under periodic reinforcement (PR) might be considered. Skinner (**16**) has reported evidence of a linear relationship between the size of the extinction ratio (the number of unreinforced responses per reinforcement) and the amount of prefeeding. Specifically, under PR, high-drive animals evince a larger extinction ratio than low-drive animals. This means that the number of unreinforced responses (extinguished or subtracted from the 'reserve') during conditioning is greater at higher drive levels than at lower, for an equal number of reinforcements. If the findings of the present study hold and drive level is not directly related to the strengthening effect of reinforcement *per se*, and if, furthermore, the 'reserve' is a function of the number of reinforcements,

then it should follow that the 'reserve' in extinction after PR is greater where conditioning has occurred under a lower as compared with a higher drive. Such an inference is readily amenable to empirical evaluation and presents a hypothesis which might be tested with profit. An additional consideration which might support the employment of PR in assessing the effect of drive variation is the fact that curves of extinction after continuous reinforcement, as Skinner (16) has pointed out, are complicated by an emotional factor which is not present during extinction after PR.

As indicated above, although drive may be ruled out as a reinforcement variable, its status as an extinction variable must still be considered. Several studies of bar-pressing in the white rat appear to establish drive as a critical factor in extinction (12, 13, 16). Thus, under increased hunger the resistance to extinction of a response previously reinforced with food appears to be greater. Yet such an inference is valid only if the unconditioned rate of responding of the animal is unaffected by an increase in drive, and there is some basis for assuming, in view of the effect of drive upon general activity, that such is not the case. Perin (12) and the others did not obtain control measures of the operant level of the organism. Hence their findings may merely reflect the influence of drive upon operant level, rather than upon resistance to extinction of the response measured. Systematic study of this hypothesis is contemplated by other investigators in this laboratory.

Summary

Bar-pressing responses of groups of white rats, maintained on a hunger rhythm and previously habituated to the serial delivery of food pellets, were conditioned in a Skinner-type apparatus after different intervals of fasting. One group received 30 regular reinforcements in deprivation sub-groups of 10–12 animals after fasting 4, 11, 23, and 47 hours, respectively. Another was given 10 reinforcements with sub-groups of similar size deprived for ½, 1, 4, and 23 hours. A final group comprising four sub-groups of 10–12 animals received a single reinforcement after 1, 11, 23, and 47 hours of fasting. Extinction to an arbitrary criterion took place under constant hunger conditions of 23 hours' deprivation in all groups. Resistance to extinction was measured in terms of number of responses, time to extinction, and rate of response.

The following conclusions were established:

1. Respecting the conditioning data there was revealed no consistent relationship between drive level and the gross time required for completion of the reinforcement process. There was, however, some evidence of a tendency for animals conditioned after 23 hours of fasting, which period is identical with the interval between daily feedings, to show a shorter 'learning time.'

2. The correlation coefficients among the several criteria in terms of which extinction data were recorded, i.e., the number of responses up to 2, 5, 8, and 10 minutes of no responding, and total number of responses for one hour of extinction, were consistently large.

3. Although resistance to extinction was evidently a function of the number of reinforcements, it was not uniformly related to the degree of hunger in conditioning at any of the three levels of reinforcement. The absence of a relationship between drive and resistance to extinction was

verified in respect to all three measures of reflex strength.

The implications of these findings in relation to the work of Finan, the earlier studies of maze learning, periodic reconditioning and resistance to extinction, and the role of drive in extinction were discussed.

(Manuscript received July 7, 1949)

References

1. Bousfield, W. A., & Elliott, M. H. The effect of fasting on the eating-behavior of rats. *J. genet. Psychol.*, 1934, **45**, 227–237.
2. Finan, J. L. Quantitative studies in motivation. I. Strength of conditioning in rats under varying degrees of hunger. *J. comp. Psychol.*, 1940, **29**, 119–134.
3. Frick, F. C., Schoenfeld, W. N., & Keller, F. S. Apparatus designed for introductory psychology at Columbia College. *Amer. J. Psychol.*, 1948, **61**, 409–414.
4. Hull, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
5. Keller, F. S., & Schoenfeld, W. N. *Principles of psychology.* (In preparation)
6. Kendler, H. H. Drive interaction: II. Experimental analysis of the role of drive in learning theory. *J. exp. Psychol.*, 1945, **35**, 188–198.
7. Koch, S., & Daniel, W. J. The effect of satiation on the behavior mediated by a habit of maximum strength. *J. exp. Psychol.*, 1945, **35**, 167–187.
8. Leeper, R. The role of motivation in learning: a study of the phenomenon of differential motivational control of the utilization of habits. *J. genet. Psychol.*, 1935, **46**, 3–40.
9. Lewin, K. Field theory and learning. In: *41st Yearbook Nat. Soc. Study Educ.* Bloomington: Public School Publishing Co., 1942.
10. Ligon, E. M. A comparative study of certain incentives in the learning of the white rat. *Comp. Psychol. Monogr.*, 1929, **6**, No. 2, 1–95.
11. Pavlov, I. P. *Conditioned reflexes.* (Tr. by G. V. Anrep.) London: Oxford Univ. Press, 1927.
12. Perin, C. T. Behavior potentiality as a joint function of the amount of training and the degree of hunger at the time of extinction. *J. exp. Psychol.*, 1942, **30**, 93–113.
13. Sackett, R. S. The effect of strength of drive at the time of extinction upon resistance to extinction in rats. *J. comp. Psychol.*, 1939, **27**, 411–431.
14. Schoenfeld, W. N., Antonitis, J. J., & Bersh, P. J. Unconditioned response rate of the white rat in a bar-pressing apparatus. *J. comp. physiol. Psychol.* (In press)
15. Siegal, P. S. A note on the construction of Skinner box pellets. *J. comp. Psychol.*, 1944, **37**, 233–234.
16. Skinner, B. F. *The behavior of organisms.* New York: D. Appleton-Century, 1938.
17. Tolman, E. C., Honzik, C. H., & Robinson, E. W. The effect of degrees of hunger upon the order of elimination of long and short blinds. *Univ. Calif. Publ. Psychol.*, 1930, **4**, No. 12, 189–202.
18. Tolman, E. C. *Purposive behavior in animals and men.* New York: Century, 1932.
19. Warner, L. H. A study of hunger behavior in the white rat by means of the obstruction method. *J. comp. Psychol.*, 1928, **8**, 273–299.
20. Williams, S. B. Resistance to extinction as a function of the number of reinforcements. *J. exp. Psychol.*, 1938, **23**, 506–522.
21. Youtz, R. E. P. Reinforcement, extinction, and spontaneous recovery in a non-Pavlovian reaction. *J. exp. Psychol.*, 1938, **22**, 305–318.

THE EFFECT OF EXTINCTION UPON RATE OF RECONDITIONING[1]

BY JAMES DEESE

Department of Psychology, The Johns Hopkins University

Introduction

Despite the fact that the problem is theoretically significant, there has been little work done upon the effects of extinction on the rate of reconditioning. The problem bears some relation to Melton's (8, 9) two-factor-theory of retroactive inhibition. It will be recalled that this theory attributes an important part of the retroactive inhibition effect in verbal learning to 'unlearning' of the original material during the learning of interpolated material. Melton cautiously avoided identifying this unlearning with either experimental extinction or with punishment. However, since subsequent investigators (1, 12) have pointed out the similarlity of unlearning to extinction, the properties of extinction in something like the retroactive inhibition experiment become important.

There are three questions which this experiment seeks to answer: (1) Is there a measurable effect upon rate of reconditioning of simple extinction? (2) Is there a measurable effect upon rate of reconditioning of extinction during the conditioning of an interpolated response, and does this extinction differ from simple extinction? (3) What is the effect of generalized extinction during conditioning of an interpolated response upon reconditioning of an original response?

Procedure

Dogs were used as subjects in a basic conditioned flexion situation described elsewhere (7). The unconditioned stimulus consisted of a single make-break d.c. shock. The intensity of the shock could be adjusted so as to produce an unconditioned response of constant magnitude. The conditioned stimuli were pure tones, 75 db *re* 10^{-16} watts/sq. cm. in intensity, produced by an audio-oscillator. The conditioned stimuli were delivered 9/5 sec. before the shock and were two sec. in duration. During conditioning and extinction, trials were administered in blocks of 20 per day on alternate days. Four unreinforced trials were randomly dispersed among each series of 20. Trials were spaced an average of 60 sec. apart, varying from 15 to 105 sec. apart in a predetermined random order.

In order to examine the effects of extinction of the original response during interpolated conditioning it was necessary to eliminate any effects of competition between the responses, either at interpolated conditioning or at reconditioning of the original response. This was accomplished by supporting the animals in such a way as to permit both responses to occur simultaneously if necessary, as had been done in previous experiments by Kellogg and his associates (5, 6). During original conditioning the unconditioned stimulus was delivered to the right rear leg, while during interpolated conditioning the unconditioned stimulus was delivered to the left rear leg.

Four groups of four animals each were used. In order to reduce the variability between subjects, all animals were subjected to two weeks of daily habituation to the experimental stock. All groups received the same treatment during original conditioning. Original conditioning consisted of an 800-cycle tone paired with a shock to the right rear leg for 120 trials spaced over six experimental periods. Following this, each group was given a different treatment for three experimental days. Finally, all groups were reconditioned for 60 trials with the 800-cycle tone paired with the shock to the right rear leg. The different experimental treatments during the interpolated period were as follows:

Group 1 (Control): This group was designed to test the effects of an interval of time, corresponding to the length of time of extinction or interpolated conditioning, between original conditioning and reconditioning.

Group 2 (Extinction): This group was subjected to 60 unreinforced presentations of the

[1] This is a portion of a dissertation submitted for the Ph.D. degree in the psychology department of Indiana University.

conditioned stimulus (800-cycle tone) between original conditioning and reconditioning. These extinction trials were given in groups of 20 trials per day.

Group 3 (Extinction during interpolated conditioning): This group was given 60 presentations of the conditioned stimulus (800-cycle tone) paired with a shock to the left rear leg between original conditioning and reconditioning. Thus extinction of the old response occurred during the acquisition of a new response.

Group 4 (Generalized extinction during interpolated conditioning): This group was given 60 presentations of a *different* conditioned stimulus (a tone of 528 cycles at 75 db) paired with a shock to the left rear leg between original conditioning and reconditioning. Thus, extinction of the old response was to a generalized stimulus. Reconditioning of the original response was to the original stimulus.

In all groups electrodes for delivering the shock were always placed on both rear legs. The greatest care was taken to give the animals no cue for a change in experimental conditions.

Latency, frequency of conditioned responses, and amplitude of conditioned responses in both legs were recorded by a pneumatic system upon kymograph paper. The experimental room, apparatus, etc., are described in detail in Kellogg, Davis, and Scott (7).

Results

Fig. 1 shows the percent conditioned responses in the right leg for all groups under all conditions. The data for the four groups have been averaged for original conditioning, since treatment was the same here, and variation in mean percent conditioned responses from group to group was very small. Table I shows the mean number of conditioned responses during reconditioning made by each group. The control group shows a much higher number of conditioned responses. As can be seen from Fig. 1, there is better than 100 percent savings in the control group over the three-day retention period. This, of course, is not remarkable, since a large number of studies testify to the high retention value of conditioned responses. However, all of the experimental groups subjected to varieties of extinction show a smaller number of conditioned responses during reconditioning. These differences

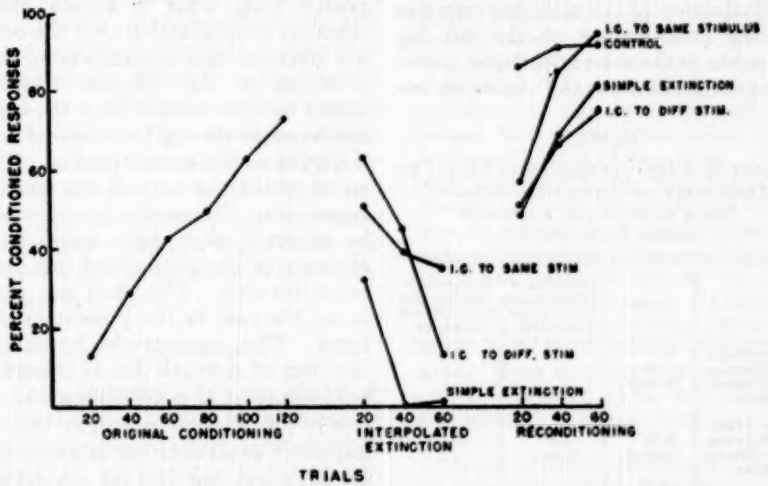

Fig. 1. Percent of conditioned responses in the right rear leg in 20 trials for all conditions. Data from all groups have been combined for original conditioning. I. C. refers to extinction during interpolated conditioning of the left rear leg.

TABLE I

MEAN NUMBER OF CONDITIONED RESPONSES DURING RECONDITIONING

	Control	Interpolated Conditioning to the Same Stimulus	Interpolated Conditioning to a Different Stimulus	Simple Extinction
Mean	53.75	38.75	47.50	39.75
Sigma	2.86	3.28	2.19	0.69

are highly significant, as can be seen from Table II.

Extinction of the right leg response during conditioning of the left leg response to the same stimulus is slower than simple extinction of the right leg response. This seems to be a result of generalization from the reinforced left leg response. This retardation in rate of extinction is reflected in greater savings upon reconditioning. More responses during reconditioning were given in the group subjected to interpolated conditioning to the same stimulus than in the group subjected to simple extinction. This difference is also highly significant (Table II).

Extinction of the right leg response during conditioning of the left leg response to the same stimulus is slower than extinction of the right leg response during conditioning of the left leg response to a different stimulus. This a a reflection of the usual failure of complete stimulus generalization. This difference in rate of extinction also appears to result in a difference in rate of reconditioning. The group subjected to interpolated conditioning to the same stimulus showed a greater savings than the group subjected to interpolated conditioning to a different stimulus. The difference in mean conditioned responses is significant.

TABLE II

VALUES OF *t* AND CORRESPONDING LEVELS OF CONFIDENCE OF DIFFERENCES BETWEEN MEAN CONDITIONED RESPONSES DURING RECONDITIONING

	Control	Interpolated Conditioning to the Same Stimulus	Interpolated Conditioning to a Different Stimulus
Interpolated conditioning to the same stimulus	2.49 (0.007)		
Interpolated conditioning to a different stimulus	8.66 (0.0005)	4.21 (0.003)	
Simple extinction	8.14 (0.0005)	3.97 (0.004)	1.02 —

Thus it appears that ease of reconditioning is inversely related to the inhibition of the response through extinction. When an interpolated response is reinforced during the extinction of the original response, the rate of extinction is retarded by a generalization from the reinforced response.

One curious paradox turns up. Why did extinction during interpolated conditioning to a different stimulus result in greater inhibition upon reconditioning than extinction during interpolated conditioning to the same stimulus? It would be expected from what is known about stimulus generalization that the original response during interpolated conditioning to the different stimulus would become weaker than the original response during interpolated conditioning to the same stimulus. This seems to be the case in the present experiment. However, it might also be expected that there would be a recovery of the generalized loss upon reconditioning. This does not seem to be the case in the present experiment. This suggests the hypothesis that loss of strength due to imperfect generalization of conditioning and recovery of response strength due to imperfect generalization of extinction are not equal, but that the generalized extinction is relatively less. This is essentially the same thing as Hull's

(4) hypothesis that the slope of the generalization of extinction curve is less than that of the generalization of conditioning.

An interesting hypothesis is suggested by the lack of a significant difference in reconditioning between the group subjected to interpolated conditioning to a different stimulus and that subjected to simple extinction. It is, of course, difficult to argue with any certainty from a negative case. However, the case is interesting enough to pursue as a possibility. This lack of a difference is even more remarkable in view of what appears to be a considerable difference in response strength between two groups on the last 20 trials of extinction. Hilgard and Humphreys (3) found that spontaneous recovery of a conditioned response was markedly affected when the response was extinguished coincidentally with the acquisition of a second response. In the case of simple extinction it is known that spontaneous recovery is considerable. However, as reported in the experiment by Hilgard and Humphreys, it is clear that spontaneous recovery in the case of such a conditioned discrimination is negligible. It then seems reasonable that there was some spontaneous recovery from simple extinction and little or none from extinction during interpolated conditioning. This might account for the lack of a difference between these groups.

Discussion

It is apparent from the above results that the amount of inhibition on reconditioning is inversely proportional to the strength of the original response at that time. Following original conditioning the strength of the original response may be lowered by extinction or raised by further conditioning. Both extinction and conditioning may appear as secondary phenomena; i.e., by way of gradients of stimulus generalization. If extinction occurs, inhibition will be the result; if conditioning occurs, facilitation will be the result.

The occurrence of simple extinction in rote verbal learning experiments appears to be largely an academic question.[3] Peak and Deese (10) have demonstrated a situation analogous to simple extinction in an experiment in verbal learning, but it is not altogether convincing. Due to the nature of the instructions generally given and the degree of sophistication of the subjects, each subject seems to provide his own reinforcement for the repetition of responses. It is not necessary for the experimenter to apply any stimulus which might be considered a reinforcer. It is apparently very difficult to isolate and remove the reinforcing agent in rote verbal learning.

Aside from the fact that the occurrence of simple extinction of verbal material may be a great experimental rarity, the operations which define it are usually clearly defined as being different from the operations which define forgetting. However, if an increasing body of evidence is to be accepted, it appears that extinction, as part of a discrimination process, leads to inhibition found in experimental studies of forgetting. It has been suggested (8, 9) that unlearning appears in retroactive inhibition in the form of elicitation of inappropriate responses through stimulus generalization. While the subject is learning the interpolated material, he may occasionally, either implicitly or overtly, tend to say words appropriate to the original material. Since these are now wrong, they are unreinforced. The inappropriateness of these responses, then, is defined

[3] It may be possible, however, that Dunlap's (2) celebrated negative practice may simply be an example of extinction through withdrawal of reinforcement from a particular mode of adjustment.

by experimental conditions analogous to the present conditioning experiment. Responses reinforced during original learning appear during interpolated learning and are not reinforced. If this were all there were to it, the matter would be relatively simple. The inhibition due to unlearning or extinction would simply be proportional to the number of unreinforced elicitations of the original material during interpolated learning. However, it is possible that other mechanisms occur which complicate the situation. The effect of reinforcement of the appropriate responses may extend to the inappropriate responses by two processes: (1) A temporal gradient of the effect of reinforcement from appropriate to inappropriate responses may exist. (2) A gradient of response generalization along some dimension other than time may exist between the reinforced responses and unreinforced responses. This creates a significant difficulty because of the lack of work upon gradients of response generalization. Lastly, as learning of the interpolated material increases, the probability of an occurrence of original material at any one place is reduced by virtue of the fact that the interpolated material comes to be dominant. This suppression of the original material may have no permanent effect at all upon the original responses, but may last only as long as the conditions for interpolated learning last.

One of the great problems in evaluating the place of extinction in the production of retroactive inhibition is the difficulty found in separating its effects from the effects of the competition of responses at the time of measurement of retention. This is further complicated by the fact that if unlearning or extinction occurs during interpolated learning it must affect the degree of competition present. This difficulty calls for the study of competition and extinction separately, which apparently cannot be accomplished in verbal learning studies.

The present paper has shown that extinction has the properties necessary to substitute for the unlearning factor in retroactive inhibition theory. Extinction results in an inhibition which can be seen in a relearning measure. Furthermore, extinction can occur during the acquisition of an alternative response, and, despite the fact that there is no competition between the two responses, extinction results in an inhibition upon reconditioning. Finally, generalized extinction during the acquisition of a second response is effective in producing inhibition upon reconditioning. Smith (11) showed that generalized extinction was effective in reducing the strength of a response. However, despite the fact that he intended his results to apply to the two-factor theory of retroactive inhibition, he did not demonstrate the effect upon reconditioning of this reduction in strength.

Finally, it should be pointed out, the author does not intend the present study to substitute for studies in retroactive inhibition. The experiment was meant to discover some of the inhibiting properties of extinction relative to retroactive inhibition. It shows the plausibility of an extinction factor operating in retroactive inhibition and suggests some of its properties. Whether or not these properties are directly relevant to retroactive inhibition in verbal learning needs to be examined by experiment.

Summary

Extinction of a conditioned flexion response results in a decrement in reconditioning as measured against a retention control. Extinction of this response during the acquisition of a second conditioned flexion response also results in a decrement in reconditioning the first response, even

though no competition exists between the responses. Extinction during the acquisition of a second response is slower than simple extinction. This is probably due to generalization from the reinforced second response. Generalized extinction (extinction to a different conditioned stimulus) of the first response during the acquisition of a second response also results in a decrement during reconditioning. The decrement in reconditioning resulting from generalized extinction during interpolated conditioning is greater than that due to extinction to the same conditioned stimulus during interpolated conditioning. This is probably due to the difference between the slopes of generalized conditioning and generalized extinction.

Extinction as a factor in the production of experimentally produced forgetting is discussed.

(Manuscript received July 24, 1949)

References

1. Bugelski, B. R. An attempt to reconcile unlearning and reproductive inhibition explanations of proactive inhibition. *J. exp. Psychol.*, 1948, **38**, 670–682.
2. Dunlap, K. *Habits, their making and unmaking.* New York: Liveright, 1932.
3. Hilgard, E. R., & Humphreys, L. G. The retention of conditioned discrimination in man. *J. gen. Psychol.*, 1938, **19**, 111–125.
4. Hull, C. L. *The principles of behavior.* New York: D. Appleton-Century, 1943.
5. Kellogg, W. N., & Walker, E. L. "Ambiguous conditioning," a phenomenon of bilateral transfer. *J. comp. Psychol.*, 1938, **26**, 63–77.
6. Kellogg, W. N., & Wolf, I. The nature of the response retained after several varieties of conditioning in the same subject. *J. exp. Psychol.*, 1939, **24**, 366–383.
7. Kellogg, W. N., Davis, R. C., & Scott, V. B. Refinement in techniques for conditioning of motor reflexes in dogs. *J. exp. Psychol.*, 1939, **24**, 318–331.
8. Melton, A. W., & Irwin, J. McQ. The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses. *Amer. J. Psychol.*, 1940, **53**, 173–203.
9. Melton, A. W., & von Lackum, W. J. Retroactive and proactive inhibition in retention: Evidence for a two factor theory of retroactive inhibition. *Amer. J. Psychol.*, 1941, **54**, 157–173.
10. Peak, H., & Deese, L. Experimental extinction of verbal material. *J. exp. Psychol.*, 1937, **20**, 244–261.
11. Smith, M. H. The effect of extinction of a generalized response on the reconditioning of the original response. *Amer. Psychologist.*, 1946, **1**, 449.
12. Underwood, B. J. Retroactive and proactive inhibition after five and forty-eight hours. *J. exp. Psychol.*, 1948, **38** 29–38.

A STUDY OF CONCEPT FORMATION AS A FUNCTION OF REINFORCEMENT AND STIMULUS GENERALIZATION [1]

ARNOLD H. BUSS

Indiana University

Introduction

Certain criteria should be met if an approach to the study of concept formation is to be useful. The measures of behavior should be objective and so far as possible reducible to physical dimensions. It should be possible to discover functional relationships between these measures and the variables determining conceptual behavior. The concepts used in dealing with such behavior should be related to those developed in the analysis of simpler behavior.

A formulation meeting these criteria might eliminate some of the confusion present in the field of conceptual behavior, the present status of which has been described as follows:

> Though there has been a growing interest of late in the psychology of conceptual activities, the many expressions of this interest in experimental and clinical research have led to no standard terminology, and to no generally acknowledged logic of investigation, experimental or otherwise, which can be presupposed in reporting research in this field. (1, p. 173.)

An important factor in this situation is the lack of rigor characterizing the traditional interpretation of conceptual behavior. This view assumes the prior existence of a concept and defines criteria for the forming, attaining, defining, discovering or acquiring of it. Experiments conducted within this framework have as the dependent variable the number of times the subject 'correctly' categorizes certain stimuli (1, 2, 6, 7, 8). However, it is not clear how laws relating 'correctness' of response to independent variables can have any generality. 'Correctness' is not a measurable property of a response; organisms do not make 'right' or 'wrong' responses, but only responses. The criteria of correctness vary in different experiments, making it probable that the experimenter's behavior, as well as the subject's, is being studied.

Another important factor in contributing to the present status of research in conceptual behavior is the relative isolation from other areas of psychology. This is probably the result of a tacit assumption that the behavior studied is qualitatively different from that studied in other areas.

A contrary assumption would be that the behavior studied is *not* qualitatively different from the behavior studied in other fields of psychology. This assumption might lead to the application of concepts from many fields of psychology, but at present certain learning concepts would seem most applicable. Learning concepts satisfy the third criterion mentioned above, for they are based upon more elementary behavior experiments. Whether the other two criteria will be met remains to be seen.

It is the purpose of this paper to attempt an application of certain learning concepts to the study of conceptual behavior and to compare the interpretation with the traditional interpretation of conceptual behavior. The merits of each interpretation may be judged by the effectiveness with which it handles the data of the following experiment.

The Problem

The purpose of the experiment was to study concept formation as a function of two kinds of independent variables: (1) the proportion of positive and negative instances in the learning

[1] Adapted from a thesis submitted in partial fulfillment of the degree, Master of Arts, Indiana University. The research was done under the supervision of Dr. W. K. Estes.

series, and (2) properties of the test stimuli. In some respects it is similar to studies done by Smoke (7, 8), who found that negative instances in the learning series increase the accuracy but not the speed of concept formation. A positive instance is defined as a stimulus which fulfills all the requirements of the concept; a negative instance does not fulfill all the requirements of the concept.

The problem may also be stated in terms of the subject's 'tending to respond' rather than 'forming a concept.' In these terms the purpose of the experiment was to discover the functional relationship between the tendency to respond in the presence of stimuli with certain properties and the number of reinforcements of the response in the presence of these stimuli.

The stimuli in the experiment were blocks with triangular top and bottom surfaces; the distance between top and bottom (height) varied. All tall blocks were positive instances, and all short blocks were negative instances. In learning terms, the stimuli to which the response was reinforced were tall, and those to which the response was nonreinforced were short. The experimenter's saying "Right" after the subject had responded was assumed to be reinforcing; saying "Wrong" after the subject responded was assumed to be nonreinforcing.

TABLE I

EXPERIMENTAL DESIGN FOR THE LEARNING SERIES

Group	Number of Subjects	Number of Tall Blocks	Number of Short Blocks
I	20	0	8
II	20	2	6
III	20	4	4
IV	20	6	2
V	20	8	0

Experimental design.—There were five groups of subjects; each group was given a different learning series but the same criterion series of blocks. The experimental design for the learning series is shown in Table I. Following the learning series, all groups were given the criterion series which had six tall, three medium and six short blocks.

In traditional terms, the subjects 'formed' or 'attained' the concept in the learning series, and how well they achieved this was measured by the number of 'correct' responses in the criterion series. In learning terms, a response in the presence of tall blocks was reinforced, and the same response in the presence of short blocks was nonreinforced in the learning series. The strength of this response was measured in the criterion series, the measures being the number of times the response was made and the latency of the response. The number of tall blocks in each learning series is not the actual number of reinforcements given; it is the number of possible reinforcements. This results from the subject's not making the appropriate response (the response that is reinforced) to every tall block in the learning series. For instance there is the possibility of reinforcing the appropriate response eight times in Group V; this becomes an actuality only if the subject makes the appropriate response eight times. If he makes this response only six times, he is reinforced only six times.

Subjects.—The subjects were 100 college students who were taking either a first- or second-year course in psychology. There were 48 females and 52 males, ranging in age from 17 to 31 years. The major field of interest for almost all subjects was not psychology.

Stimuli.—The upper and lower surfaces of each block were approximately right triangular. Each block was placed so that the apex faced the experimenter and the base faced the subject. The heights of the blocks were as follows: tall, one in.; medium, $\frac{11}{16}$ in.; short, $\frac{1}{2}$ in. In order to make the task difficult enough to obtain differences in performance, the surface area and color of the blocks were varied. The dimensions of the top surfaces of the blocks were as follows: large area, $2.2 \times 1.6 \times 1.5$ in. = 1.20 sq. in.; small area, $1.5 \times 1.1 \times 1.0$ in. = 0.55 sq. in. Each block was painted a solid color. Black, brown, green, orange and yellow were used for the learning series; blue, gray, red, varnish and white were used for the criterion series.

Apparatus.—*S* was separated from *E* by a large wooden board which enabled *S* to see only two telegraph keys and an opening covered by a shutter, as shown in Figs. 1 and 2. The left key was marked 'VEC' and the right key 'not VEC.' When both keys were depressed, raising of the

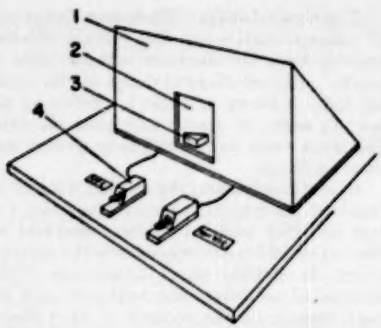

FIG. 1. Apparatus as seen from the subject's side. 1. wooden barrier, 2. opening, 3. block, 4. telegraph key.

shutter closed a circuit which activated the chronoscope. A relay kept the chronoscope going after the shutter was down. The chronoscope was stopped when release of either key broke the circuit. Release of the 'not VEC' key resulted in a tiny flash of light which only *E* could see. *E* knew that the 'VEC' key had been released when no flash occurred. The shutter was operated manually by *E*.

Procedure.—*S* was handed the following typewritten instructions:

> You will be shown a series of blocks, one at at time. Some of the blocks are VEC and some are not VEC. You will indicate whether you think a block is VEC or not VEC by lifting your finger from the appropriate key. If you think the block is VEC, lift your finger from the key marked 'VEC'; if you think the block is not VEC, lift your finger from the key marked 'not VEC.' The sequence will be as follows: Depress both keys. The shutter will be raised, and you will see a block. Indicate by lifting your finger *from one key* which you think the block is, VEC or not VEC. Then the shutter will drop, and you will be told whether your choice was right or wrong. As soon as the shutter drops, again depress both keys, and you are ready for the next block.

After the learning series *S* was handed the following instructions:

> The same conditions will hold in the next series, except that you will not be told whether you are right or wrong.

The exposure time for each block was five sec.; at the end of this time the shutter was closed whether or not *S* had responded. If *S* responded before five sec., the shutter was held open until the full exposure time had elapsed. As soon as the shutter was closed, *E* recorded the response (VEC or not VEC), whether it was 'correct' or 'incorrect,' and the latency. Then the block was replaced on the platform behind the shutter by the next block in the series, and the shutter was opened. The average time between closing the shutter for one block and opening it for the next was approximately 10 sec. The average time between learning and criterion series was approximately one min. Records were kept for both learning and criterion series.

The order of presentation of the blocks varied in the learning series, but there was a fixed order of presentation in the criterion series. In the following list of criterion series blocks 'tall' and 'short' refer to height, and 'large' and 'small' refer to area: 1. tall large gray; 2. short large red; 3. tall small varnish; 4. short small white; 5. medium small blue; 6. short large gray; 7. short small red; 8. tall large blue; 9. medium small varnish; 10. tall small white; 11. short large blue; 12. medium small gray; 13. tall large white; 14. short small varnish; 15. tall small red.

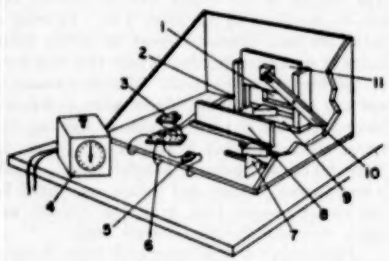

FIG. 2. Apparatus as seen from the experimenter's side. 1. opening, 2. contact, 3. relay, 4. chronometer, 5. signal light, 6. wire to telegraph key, 7. block, 8. barrier, 9. platform, 10. lever, 11. platform.

RESULTS

The dependent variable in the traditional interpretation of conceptual behavior is the number of 'correct' responses. A comparison of the various experimental groups with respect to this variable is presented in Table II. Since only the responses to the six tall blocks and the six short blocks in the criterion series were counted as 'correct' or 'incorrect,' the maximum number 'correct' per subject was 12. The means of the total number of 'correct' responses varies

TABLE II

MEAN NUMBER OF 'CORRECT' RESPONSES AS A FUNCTION OF THE NUMBER OF POSITIVE INSTANCES IN THE LEARNING SERIES

	Number of Positive Instances				
	0	2	4	6	8
Tall Blocks	2.00	3.40	3.80	4.25	5.20
Short Blocks	5.35	4.10	3.95	3.40	1.75
Total	7.35	7.50	7.75	7.65	6.95

little between groups; analysis of variance shows that the error variance exceeds the variance between groups. The lack of significant differences between groups suggests that the independent variable has little effect upon the number of 'correct' responses.

Fig. 3 shows the breakdown of total number of 'correct' responses into those for tall and short blocks. Groups presented with more tall blocks in the learning series had more of these 'correct' in the criterion series. Groups having more short blocks in the learning series had more of these 'correct' in the criterion series. Each group had approximately the same number 'correct' because there were as many tall blocks as short blocks in the criterion series. Since subjects seeing tall blocks in the learning series tended to categorize them correctly in the criterion series, their total number 'correct' could be raised by adding tall blocks to the criterion series. The same is true for subjects seeing short blocks in the learning series. Thus, the total number of 'correct' responses is a function of the blocks constituting the criterion series as well as of positive and of negative instances in the learning series.

In learning terms, the dependent variable is the tendency to make the VEC response, as measured by the number of VEC responses. The data on the number of VEC responses are presented in Table III. The total number of VEC responses made by the 20 subjects in each group was divided by the number of subjects in

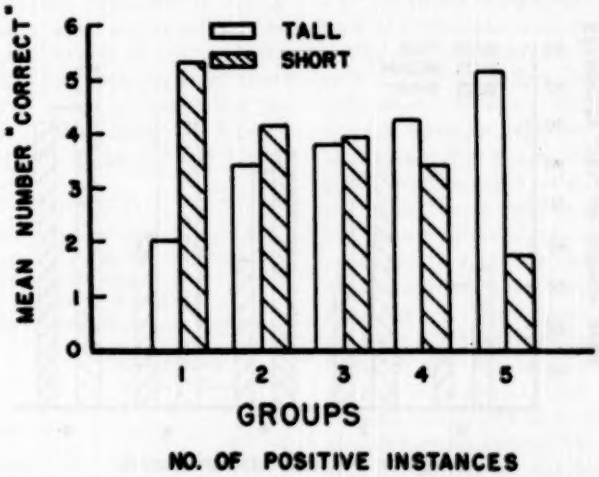

FIG. 3. Mean number of 'correct' responses as a function of the number of positive instances in the learning series.

TABLE III

MEAN NUMBER OF VEC RESPONSES PER BLOCK AS A FUNCTION OF NUMBER OF POSSIBLE REINFORCEMENTS IN THE LEARNING SERIES

	Number of Possible Reinforcements				
	0	2	4	6	8
Tall Blocks	.33	.57	.63	.71	.87
Medium Blocks	.30	.57	.57	.73	.75
Short Blocks	.11	.32	.36	.43	.71
Total	.24	.47	.50	.60	.78

the group, and the resulting figure was divided by the number of blocks responded to; this yielded the mean number of VEC responses per block. For instance, the subjects in Group I made a total of 40 VEC responses in the presence of the six tall blocks in the criterion series. This figure was divided by the number of subjects (20), and the result was divided by the number of tall blocks (6) to yield a mean number of VEC responses per block of .33.

These data indicate that the number of VEC responses increases as a function of the number of possible reinforcements in the learning series. Analysis of variance on total number of responses yields an F of 24.9, significant at the .01 level of confidence. The correlation ratio (eta) of .72 suggests a strong relationship between the independent and dependent variables. The F's for VEC responses in the presence of tall, medium and short blocks were 8.56, 7.05, and 13.40, respectively, all significant at the .01 level of confidence; the correlation ratios were .51, .47 and .60, respectively.

It was mentioned earlier that the number of tall blocks in the learning series is the number of possible reinforcements and that this differs from the actual number of reinforcements. The average actual numbers of reinforcements were 0, 1.05, 2.55, 4.00, and 6.35 for Groups I, II, III, IV, and V, respectively. Each of these values except the first is lower than the number of possible reinforcements for each group. This indicates that the

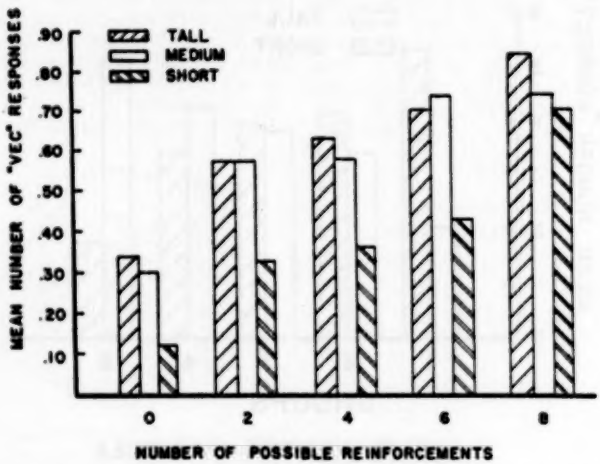

FIG. 4. Mean number of VEC responses per block as a function of possible reinforcements in the learning series.

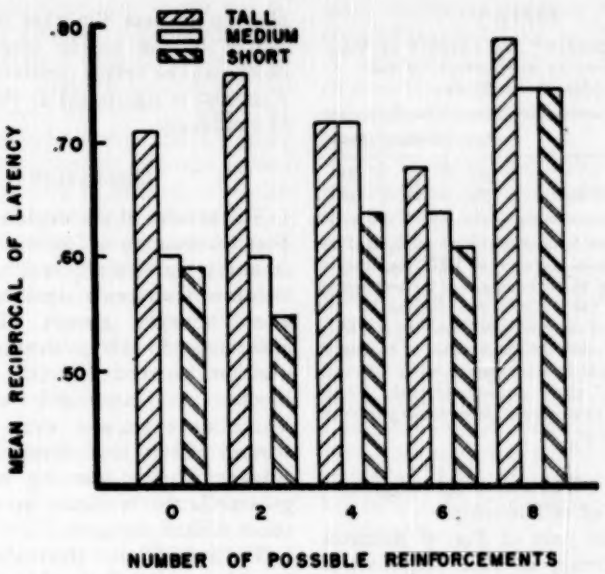

FIG. 5. Mean reciprocal of latency as a function of possible reinforcements of VEC responses in the learning series.

rate at which the number of VEC responses increases as a function of the number of reinforcements is greater than that indicated by Fig. 4.

Latencies of VEC responses may be regarded as measures of the tendency to respond. The reciprocals of these were used in order to normalize the distribution. The data on latency are presented in Table IV and Fig. 5.

TABLE IV

MEAN RECIPROCAL OF LATENCY OF VEC RESPONSES AS A FUNCTION OF POSSIBLE REINFORCEMENTS IN THE LEARNING SERIES

	Number of Possible Reinforcements				
	0	2	4	6	8
Tall Blocks	.71	.76	.72	.68	.79
Medium Blocks	.60	.60	.72	.64	.75
Short Blocks	.58	.55	.64	.61	.75
Total	.63	.65	.70	.65	.77

There seems to be a definite relationship between the latencies and the number of reinforcements; the F of 7.31 for totals is significant at the .01 level of confidence, and eta is .48.

The VEC response is made not only to tall blocks but also to medium and short blocks. Evidence from certain learning experiments (3, 4, 5, 9) suggests that the further away on a given continuum a stimulus is from the stimulus to which the response has been reinforced, the smaller is the tendency to make the response. The data on the tendency to make the VEC response in the presence of tall, medium and short blocks are shown in Table V and Fig. 6. The upper part of Fig. 6 indicates that the number of VEC responses is an inverse function of the distance of the blocks from the tall blocks on the height continuum. Analysis of variance yields an F of 9.7, significant at the .01 level of

TABLE V

MEAN FREQUENCY* AND LATENCY OF VEC RESPONSES AS A FUNCTION OF THE HEIGHT OF BLOCKS

	Tall	Medium	Short
Frequency	.64	.59	.38
Reciprocal of Latency	.73	.66	.58

*The mean frequencies were computed as follows: the total number of VEC response in the presence of all of the blocks of a given height made by all the subjects was divided by the total number of subjects (100), and the resulting figure was divided by the number of blocks of the given height in the criterion series (six tall, six short, and three medium blocks). These computations yielded mean frequencies per block of a given height.

confidence. The eta of .46 suggests a fair degree of relationship.

The lower part of Fig. 6 indicates that the latency of VEC responses is also an inverse function of the distances of the blocks from the tall blocks on the height continuum. The F of 6.98 is significant at the .01 level of confidence.

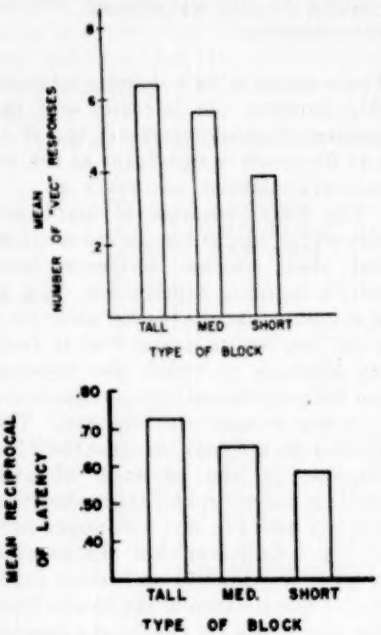

FIG. 6. Mean frequency as latency of VEC responses as a function of the height of blocks.

DISCUSSION

The results of the experiment show that the number of 'correct' responses is not a particularly useful measure because it conceals significant differences between groups. That such differences do exist is shown when the number 'correct' for tall and short blocks are presented separately. These data suggest that the more stimuli of one kind presented to the subject in the learning series, the greater is the tendency to categorize these stimuli correctly.

The results also show that a 'concept' may be 'formed' with either positive or negative instances in the learning series, or both. This conclusion holds only for experiments in which there are equal numbers of positive and negative instances in the criterion series. If the criterion series consisted entirely of positive instances, negative instances would not help 'form the concept'; the same is true for positive instances when the criterion series consists entirely of negative instances. These findings suggest that the relative number of positive and negative instances is a factor in conceptual behavior. However, the measures traditionally used do not furnish information concerning the influence of this factor because they are primarily a function of the kind of stimuli used in the test series.

Perhaps the results of the experiment will become more meaningful when expressed in learning terms. As yet no account of concept formation in terms of present-day learning theory has appeared in the literature.

Such an interpretation might be developed as follows. The 'concept' is a class of stimuli with certain properties; in the present experiment the class of stimuli was tall blocks. The measure of behavior is the tendency to make a particular response to one or another class of stimuli. In this experiment the response was releasing the VEC key in the presence of tall, medium and short blocks, measures of the tendency to respond being frequency and latency. The tendency to make the response should be a function of many of the same variables that control simpler kinds of learning, one of the most important being the number of reinforcements. The greater the number of reinforcements, the greater is the tendency to make the VEC response. This relationship is reflected in Table III and in the eta of .72 between number of possible reinforcements and frequency of VEC responses by groups.

The various stimulus properties comprise a second variable affecting the tendency to make a VEC response. It is expected that this tendency to respond is greater in the presence of stimuli closer on the height continuum to tall blocks than it is in the presence of stimuli further away on the height continuum. The results for both frequency and latency are in the direction predicted, as illustrated in Fig. 6. These gradients are similar to stimulus generalization gradients such as have been demonstrated in simple conditioning experiments (3, 5).

Another variable affecting the tendency to respond VEC is nonreinforcement of the VEC response in the presence of stimuli other than tall blocks. This is expected to result in at least a partial extinction of the VEC response in the presence of non-tall blocks. Where such nonreinforcement occurs the gradient in the tendency to respond VEC in the presence of different stimuli on the height continuum should be steeper than it would be if it were due only to stimulus generalization.

The learning series for Group V consisted of eight tall blocks and no short blocks; thus there was no possibility of nonreinforcement of the VEC response in the presence of short blocks. Any differences between the tendency to respond VEC in the presence of short blocks and this tendency in the presence of tall blocks may be accounted for in terms of a primary stimulus generalization gradient. The learning series for Group I consisted of eight short blocks and no tall blocks, making possible eight nonreinforcements of the VEC response in the presence of short blocks. It is expected that these nonreinforcements should make the gradient in the tendency to make the VEC response to different stimuli on the height continuum steeper than could be accounted for solely by the operation of a primary stimulus generalization gradient.

A measure for comparing the steepness of the gradients for the different groups is the ratio of the frequency of VEC responses in the presence of short blocks to the frequency of VEC responses in the presence of tall blocks. This ratio is approximately .8 for Group V and .3 for Group I. Since only stimulus generalization was operating in Group V, it seems reasonable to attribute the steeper gradient in Group I to the additional factor of nonreinforcement of the VEC response in the presence of stimuli other than tall blocks. Stated in other terms, nonreinforcement of the VEC response in the presence of short blocks made for better discrimination between tall and short blocks.

The ratio mentioned in the preced-

ing paragraph may be calculated from the data given in Tables III and IV. The data in Table III show that for Groups II, III, and IV the ratios fall between the values mentioned above for Groups I and V. Table IV shows that the ratios for the latency data follow a pattern similar to that of the frequency data. These differences in the tendency to respond are in the direction predicted by a theory of discrimination learning such as that of Spence (9).

The interpretation of the experiment outlined above uses a consistent terminology, one in common use in a field of psychology other than conceptual behavior. The proposed measures of behavior appear adequate to handle the data in a standard concept formation experiment. Since the interpretation promotes meaningful exposition of the data, the application of other concepts as well as methods, techniques and terminology from the field of learning would seem a promising solution to the present-day problems in the study of conceptual behavior.

Summary

An experiment on concept formation was performed in order to compare the effectiveness with which two interpretations would handle the data. In traditional terms the purpose of the experiment was to study concept formation as a function of (1) the proportion of positive and negative instances in the learning series; and (2) properties of the test stimuli. In learning terms the purpose was to discover the functional relationship between the tendency to make a particular response to stimuli with certain properties and the number of reinforcements of the response in the presence of these stimuli. The stimuli were blocks which varied in height, color and surface area; the top and bottom surfaces were triangular. There were five learning series, each having a different proportion of tall blocks (positive instances) and short blocks (negative instances). The subjects, 100 college students of both sexes, were divided into five groups of 20 each. Each group had one of the five learning series followed by a criterion series which was the same for all groups. The blocks were presented singly for five seconds, and the subject released one of two telegraph keys to indicate his choice. The subject was told "Right" or "Wrong" after each response in the learning series but not in the criterion series. The number 'correct' and the frequency and latency of responses were recorded.

The results indicated that the number 'correct' is as much a function of the kind of blocks in the criterion series as of the proportion of positive and negative instances in the learning series. The frequency and latency measures were found to be functionally related to (1) reinforcement of a particular response in the presence of tall blocks; (2) nonreinforcement of this response in the presence of stimuli other than tall blocks; and (3) certain stimulus properties. An account of the experiment in learning terms was proposed, and certain advantages that might result from further applications of learning theory to the study of conceptual behavior were mentioned.

(Manuscript received July 5, 1949)

References

1. Hanfmann, E., & Kasanin, J. A method for the study of concept formation. *J. Psychol.*, 1937, 3, 521–540.

2. Heidbreder, E. The attainment of concepts: I. Terminology and methodology. II. The problem. *J. gen. Psychol.*, 1946, 35, 173–191, 191–225.
3. Hovland, C. I. The generalization of conditioned responses. I. The sensory generalization of conditioned responses with varying frequencies of tone. *J. gen. Psychol.*, 1937, 17, 125–148.
4. Hull, C. L. The problem of stimulus equivalence in behavior theory. *Psychol. Rev.*, 1939, 46, 9–30.
5. Hull, C. L. *Principles of behavior.* New York: Appleton-Century-Crofts, 1943.
6. Reed, H. B. The learning and retention of concepts. II. The influence of length of series. III. The origin of concepts. *J. exp. Psychol.*, 1946, 36, 71–87, 166–179.
7. Smoke, K. L. An objective study of concept formation. *Psychol. Monog.*, 1932, 42, No. 4.
8. Smoke, K. L. Negative instances in concept learning. *J. exp. Psychol.*, 1933, 16, 583–588.
9. Spence, K. W. The differential response in animals to stimuli varying in a single dimension. *Psychol. Rev.*, 1937, 44, 430–444.

THE LEARNING AND RETENTION OF CONCEPTS. V. THE INFLUENCE OF FORM OF PRESENTATION

BY HOMER B. REED

Fort Hays Kansas State College

In previous studies (1, 2, 3) we investigated how attitude or set influences the economy of learning and retaining concepts, how the amount of learning effort changes in relation to the length of the series from which the concepts are derived, what the processes are by which concepts originate, and how the learning effort and its accuracy are influenced by complicating the stimuli from which the concepts are to be derived. In this paper we investigate how the economy of learning and retaining concepts is influenced by the form of presentation in relation to the length of the series. More specifically, there are four problems: 1. What is the comparative economy of serial versus simultaneous presentation in the formation and retention of concepts? 2. How is the economy changed by varying the length of the series? 3. How is the frequency of acquiring accurate or consistent concepts influenced by the form of presentation? and 4. How is this frequency changed by varying the length of the series?

Method and Materials

The materials were the same as those described in (1). To describe them again briefly, there were 42 cards, seven for each of six nonsense syllables in the serial presentation. In the front of each card were printed four disconnected English words, one of which, called the critical word, belonged to a logical group designated by a nonsense syllable, which was printed on the back. The subject's task was to learn the name of each card by the prompting method and to figure out the meaning of each syllable by relating it to the critical words to which it belonged. These cards were shown at the rate of one every seven seconds until the subject responded correctly to each of the cards in a single trial. Learning effort was measured by the number of promptings per concept or syllable required to reach the first errorless trial.

In the simultaneous presentation, exactly the same words were used, but instead of four words being printed on a single card, they were printed in a line on typing paper, and all the lines relating to one syllable appeared on one page headed by the proper syllable. All the words related to the concept in question were before the subject in a single field of perception. His task was to find the group of words that belonged to the syllable and give its meaning. The subject was allowed to study until he thought he knew what the syllables meant and told the experimenter the correct meaning of at least one of the syllables, after which the typed sheets were removed and serial presentation was given as a test of the correctness of the subject's concepts. To make the results comparable to those from the group which used only serial presentation, the presentation of the cards was continued until the subject responded correctly to each of the cards in a single trial. After a subject reached the criterion, he was asked to tell the meaning that he had worked out for each syllable and by what process he had formed it. A record was made of these reports, of any pertinent remarks that he made during learning or study, of the time used for study, of the number of promptings required for the serial presentation, of the erroneous responses made to the cards, and of the total time between the start of the learning and the reaching of the criterion.

A retention test was given to each subject one week after the learning. In this test only serial presentation was used.

To provide a common measure for both forms of presentation, the number of seconds used for study during the simultaneous presentation was divided by seven, the number of seconds used to present one card or give one prompting. This made it possible to measure the total learning effort in terms of the equivalent of the number of promptings.

A total of 118 college students formed into equivalent groups by means of the Henmon-Nelson Test of Mental Ability for College Students was used as subjects. They were divided in six groups as shown in Table I.

For this study, Groups 6, 4, and 7 were the control groups, and Groups 13, 12, and 14 were

TABLE I

SERIAL PRESENTATION

Group	Number of Subjects	Lengths of Series
6	13	24
4	19	42
7	21	60

SIMULTANEOUS PRESENTATION

Group	Number of Subjects	Lengths of Series
13	20	24
12	24	42
14	21	60

the experimental groups. The results for the control groups are taken from an earlier paper (2).

To give a more exact idea of the form of presentation, the materials and the procedure, we present here the materials for two of the syllables and a copy of the directions read to the subject:

Find the meaning of each of the nonsense syllables.

KUN:	1. horn	leaf
	2. fame	tiger
	3. line	people
	4. uncle	fried
	5. crowd	sail
	6. horse	circle
	7. carrying	cow
VOR:	8. brook	leave
	9. lover	brow
	10. breaking	idle
	11. broken	darling
	12. barrel	noisy
	13. banana	haste
	14. believe	cigar
		monkey
		ought
		elephant
		pear
		deer
		paid
		died
		claim
		flower
		honey
		load
		hurried
		dear
		owe
		debt
		saucer
		sound
		sheep
		string
		letter
		rule
		precious
		point
		bread
		pearl
		sweetheart
		minutes
		love

Directions Number Five

Learning Concepts, Simultaneous Presentation.

This is an experiment in learning concepts A concept, you know, is a word, or idea that stands for any one of a group of things. Thus, the word chair, bird, or stone stands for no particular chair, bird, or stone, but for any *one* of a *group* of chairs, birds, or stones.

I am going to show you two typewritten pages containing six groups of words, each of which is headed by a nonsense syllable such as, jok, bik, or bex. Each nonsense syllable is a concept. Look carefully at all the words in each group and find, if you can, what the syllable stands for. As soon as you have found the meaning of all the syllables, hand me the typewritten pages, then I shall show you cards each of which contains one of the lines which you are to study. You are to tell me the nonsense syllable that belongs to that line. I shall always prompt you when you fail to name the card within three seconds. When I have prompted you, repeat the syllable aloud after me so that I can be sure that you understand it. Your work will be finished as soon as you can name each card without any help. Now will you please answer these questions:

1. What is this an experiment in?
2. What is a concept?
3. In this experiment is each nonsense syllable a concept?
4. What are you to do?

After clear and correct answers were obtained to the above questions the experiment was begun.

RESULTS

Some guesses that were made about the character of the results in advance of the experiment were that simultaneous presentation would be more economical than serial because it would be less of a strain on one's memory to see all the facts relating to one concept than to get the same facts at successive intervals over a longer period of time, that the shorter the series the quicker the learning would be since the chances of forming the right concept from a small number of instances would be fairly high, and that the accuracy of

TABLE II

MEAN NUMBER OF PROMPTINGS PER CONCEPT FOR SERIAL AND SIMULTANEOUS PRESENTATION

Form of Presentation	Group	N	Mean Number of Promptings Per Concept				Percent Retention
			Learning Mean	S.D.	Relearning Mean	S.D.	
Simultaneous	4	19	30.70	13.40	1.35	1.76	97.8
Serial	12	24	30.35	32.25	.54	.69	92.0
	Difference		.35		.81		

the concepts would be proportional to the length of the series since a hypothesis confirmed by a large number of facts has a higher validity than one confirmed by a small number. After the experimental results were obtained, we discovered that the first two guesses were pretty bad but that the third one was good.

1. The comparative economy of serial and simultaneous presentation.

The first experiment on simultaneous presentation was made on Group 12 with 24 subjects and 42 cards. The results are compared with those for serial presentation obtained from Group 4, and are given in Table II.

This table indicates that from the standpoint of economy the two methods are equal for learning concepts, but the simultaneous method appears to be slightly superior to the serial for retention. In a previous study (1) we discovered that the amount of effort required to form consistent or correct concepts was very much less than for inconsistent or incorrect concepts. When we studied the percentage of correct concepts in the simultaneous presentation, we discovered that 99 percent of them were correct for Group 12 but that only 86 percent of them were correct for Group 4. It would therefore be more accurate to compare the learning effort of Group 12 with the effort which Group 4 required to form the correct concepts. This is done in Table III.

We now see a greater difference in favor of serial presentation but it is not large enough to be significant. For practical purposes the two methods for the length of series used are equal.

2. The comparative economy of simultaneous and serial presentation in relation to length of series.

TABLE III

MEAN NUMBER OF PROMPTINGS PER CORRECT CONCEPT FOR SIMULTANEOUS AND SERIAL PRESENTATION

Form of Presentation	Group	N	Learning		Relearning		Percent Retention
			Mean	S.D.	Mean	S.D.	
Simultaneous	12	24	30.35	32.25	.54	.69	98
Serial	4	19	25.40	10.53	1.04	1.14	92
	Difference		4.95	7.15	−.50	.30	

The next question is whether the economy of the two methods changes when the length of the series is varied. To solve this problem, two other groups, Groups 13 and 14, were organized to learn concepts from series of 24 and of 64 lines or cards respectively. Except for length of series, the method and the procedure were the same as for Group 12. Their results were compared with the results from Groups 6 and 7 which had used serial presentation. The comparative findings are given in Table IV and Fig. 1.

For the short series, 24 cards, the number of promptings required to form the correct concepts is significantly less for serial than for simultaneous presentation. Within the limits of the experiment the difference in the economy of the two methods approaches zero as the length of the series is increased.

From the standpoint of retention the two methods are practically equal.

The relation of the amount of work to the length of the series for the two series is not the same for the two forms of presentation. For the simultaneous presentation the amount of work required to form the correct concepts decreases sharply from 24 cards to 42 cards and after that remains about the same. For the serial presentation the amount of work to form correct concepts varies directly as the length of the series. These results indicate that in simultaneous presentation a certain minimum length of series is necessary to form correct concepts with an economical amount of effort. If the length of the series is below this minimum, the amount of work is greatly increased. Two questions may be raised about Table IV: 1. Why in case of simultaneous presentation do 24 cards require more work than 42 cards? and 2. Why is this not true for serial as well as for simultaneous presentation? In our previous studies, we found that the introduction of conflicting concepts increased the amount of work necessary to form a correct concept, and also the probability of not getting the correct concept. An examination of our learning material shows that the pro-

TABLE IV

MEAN NUMBER OF PROMPTINGS FOR LEARNING AND RELEARNING CORRECT CONCEPTS BY METHODS OF SERIAL AND SIMULTANEOUS PRESENTATION IN RELATION TO LENGTH OF SERIES

Group No.	Length of Series	N	Method	Learning		Relearning		Percent Retention
				Mean	S.D.	Mean	S.D.	
13	24	20	Simultaneous	42.50	32.05	.95	.93	95
6	24	13	Serial	18.90	6.31	1.24	1.28	93
	Difference			23.60	7.57	−.29	.39	
12	42	24	Simultaneous	30.35	32.25	.54	.69	98
4	42	19	Serial	25.40	10.53	1.04	1.14	92
	Difference			4.95	7.15	−.50	.30	
14	60	21	Simultaneous	31.05	22.70	.68	.94	98
7	60	21	Serial	29.35	19.20	.51	1.52	98
	Difference			1.70	6.95	.17	.23	

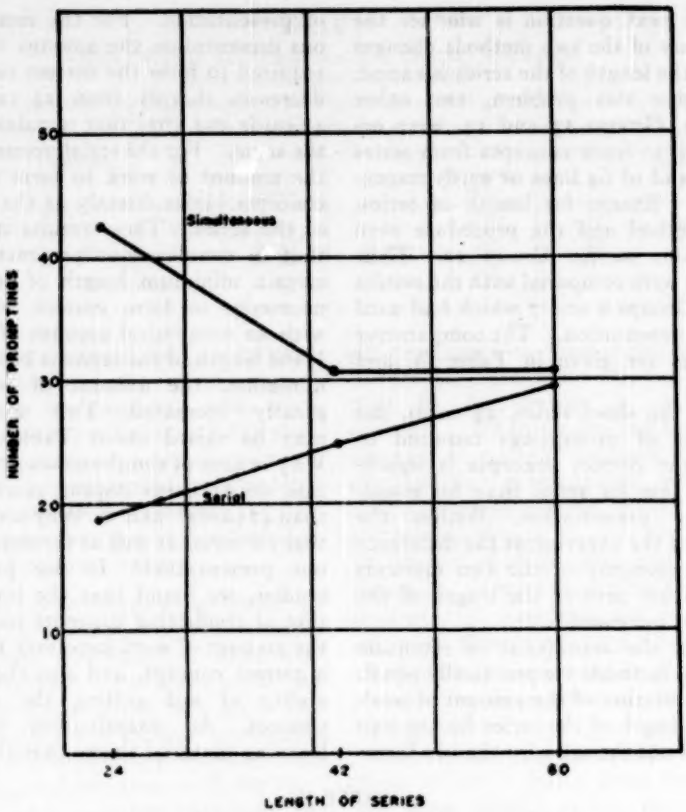

FIG. 1. Number of promptings required to form concepts in relation to form of presentation and length of series.

portion of supporting evidence for a conflicting concept is greater in the short than in the long series. For example, in connection with the syllable, *BEP*, four of the lines each contain the name of an animal and three of them contain a verb. If the word, pasture, is thought of as a verb, as it might be, then there is just as much evidence for saying that *BEP* means a verb as for saying that it means a vegetable. However, when seven lines are presented there is no consistent evidence for the concept, verb, and this hypothesis is therefore eliminated. According to this illustration, the greater amount of work for the short series is not due to its length but rather to conflict, or the fact that the evidence in favor of the designated correct concept is not so clear cut as in the longer series.

The fact that the ratio of amount of work to length of series is not so consistent in simultaneous as in serial presentation is due probably to the method of learning or to a difference in the subject's interpretation of his task. In simultaneous presentation the subject concentrated primarily

TABLE V

THE PERCENTAGE OF CORRECT CONCEPTS IN RELATION TO FORM OF PRESENTATION

Group	N	Form of Presentation	Percent Correct Concepts
12	24	Simultaneous	99
4	19	Serial	86
		Difference	13
		$\sigma_{Diff.}$	8.15
		t	1.59

upon working out the meaning of the syllables and only secondarily upon learning the name of the syllable that goes with each line. In the serial presentation he concentrated upon learning the names of the cards rather than upon working out their meanings. Since the evidence for the designated meanings was less clear cut in the short than in the long series, the amount of work required to reach the goal was greater.

3. The relation of the form of presentation to the frequency of the correct concept.

We have already stated that our initial hypothesis was that simultaneout presentation would favor the accuracy of the concepts. Our results support this idea. Table V gives the results for the series of 42 cards.

We see that simultaneous presentation yields 13 percent higher accuracy than the serial. Although the difference is not highly significant, it is close to the 10 percent level. The reason for the difference is rather obvious. When one can see both the supporting and the non-supporting evidence for a certain hypothesis before him in a single field of perception, he can draw a more accurate conclusion than when the evidence is spread over different fields.

4. The accuracy of concepts for simultaneous and serial presentation in relation to the length of the series.

How does the superiority of simultaneous over serial presentation for accuracy of concepts change in relation to the length of the series? To answer this question we made calculations similar to those in Table IV for the groups working on the 24-card and the 60-card series. The results are given in Table VI and in Fig. 2.

We see a large and significant difference in favor of simultaneous presentation in 24-card series and that this

TABLE VI

COMPARATIVE EFFICIENCY OF SIMULTANEOUS AND SERIAL FORMS OF PRESENTATION FOR ACCURACY OF CONCEPTS IN RELATION TO LENGTH OF SERIES

Group	N	Form of Presentation	Length of Series	Percent Correct	$\sigma_{Diff.}$
13	20	Simultaneous	24	92	
6	13	Serial	24	42	
			Difference	50	14.87
12	24	Simultaneous	42	99	
4	19	Serial	42	86	
			Difference	13	8.15
14	21	Simultaneous	60	98	
7	21	Serial	60	94	
			Difference	4	6.00

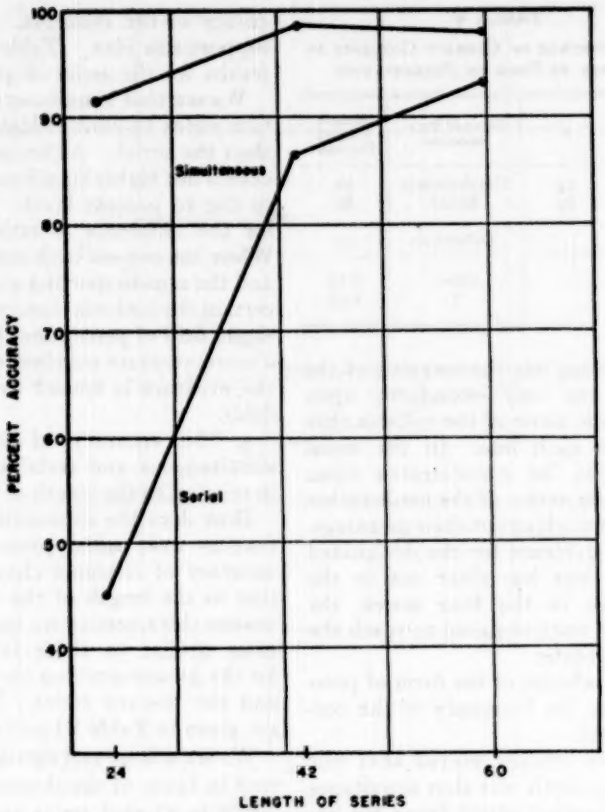

FIG. 2. Accuracy of concepts in relation to form of presentation and length of series.

difference rapidly approaches zero as the length of the series is increased to 60 cards. The difference is 50 percent for 24-card series, 13 percent for the 42-card series, and only 4 percent for the 60-card series. This indicates that for a small amount of evidence the accuracy of concepts is primarily a function of the form of presentation but as the amount of evidence is increased the form of presentation becomes insignificant and the amount of evidence becomes the principal determining factor.

SUMMARY

The problem of this investigation was to find out to what extent the learning and retention of concepts was a function of the form of presentation in relation to the length of series. The materials used were cards each of which had four unrelated words on the front and a nonsense syllable on the back. One of the words belonged to a category for which the nonsense syllable was a symbol. The subject's task was to learn the name of the card and its meaning. In the

serial presentation the cards were shown at the rate of one every seven sec. and the subject learned them by the prompting method. In the simultaneous presentation the lines of words instead of appearing on separate cards were typed on one page headed by the proper nonsense syllable. The learning effort was measured by the number of promptings per concept required to name all the cards correctly in one trial. A concept was defined as a word or idea that stands for any one of a group of things.

The conclusions reached may be stated as follows:

1. The economy of the two methods of presentation depended upon the length of the series, but for each length investigated—24 cards, 42 cards, and 60 cards—the simultaneous method required more effort than the serial. It was significantly larger for the short series, but the difference between the two methods rapidly approached zero as the length of the series was increased.

2. The ratio of the amount of work required to reach the criterion to the length of the series was different for the two modes of presentation. For the serial presentation the amount of work varied directly as the length of the series, but for the simultaneous the amount of work per concept was much greater for short series than for long series. After a certain length of series the amount of work appeared not to change with the length of the series.

3. The effect of the form of presentation upon the accuracy of the concepts depended upon the lengths of the series. For short series the percentage of correct concepts was less than half as great as for long series, but this difference rapidly approached zero as the length of the series was increased.

(Manuscript received August 2, 1949)

References

1. Reed, H. B. *Factors influencing the learning and retention of concepts. J. exp. Psychol.*, 1946, 36, 71–87.
2. Reed, H. B. *The learning and retention of concepts.* II. *The influence of length of series;* III. *The origin of concepts. J. exp. Psychol.*, 1946, 36, 166–179.
3. Reed, H. B. *The learning and retention of concepts.* IV. *The influence of the complexity of the stimuli. J. exp. Psychol.*, 1946, 36, 252–261.

SENSORY CONDITIONING MEASURED BY THE FACILITATION OF AUDITORY ACUITY *

BY W. J. BROGDEN

University of Wisconsin

INTRODUCTION

Since the original study of sensory pre-conditioning was made on dogs by Brogden (1), two different procedures for producing the phenomenon in human beings have been reported. Karn (6) used a procedure very similar to that of Brogden's original experiment on dogs. Fifty trials of a buzzer and light in combination for two sec. were presented to the experimental group in a single session. Finger withdrawal was conditioned to the sound of the buzzer in a second experimental session. Once the criterion of conditioning was met, the light was presented 10 times. Finger withdrawal to the light occurred with greater frequency in subjects of the experimental group than in subjects of a control group that were given the same treatment except for the 50 trials of buzzer and light in combination. Brogden (2) used a reaction time situation for demonstrating sensory pre-conditioning in human subjects. *S*s of the experimental group were presented with 10 trials of a tone and light in combination for two sec. under instruction to wait in the experimental room while *E* adjusted his apparatus. They were then instructed to press a telegraph key to the light. After 30 reaction time trials were given, the tone was presented 10 times. These *S*s responded with greater frequency to the tone than did the *S*s of two control groups, one of which was given no pre-treatment and the other of which was given 10 trials of the tone alone. Chernikoff and Brogden (3) altered the procedure of the above experiment by changing the source of the tone from the cabinet in which the light was located to an indefinite localization provided by a pair of loud speakers beneath the table that were equidistant from *S*. This altered procedure was more efficient in producing sensory pre-conditioning in that an *N* of 10 for each group was sufficient to demonstrate statistically significant differences between the experimental and control groups. With the unaltered procedure, an *N* of 42 for each group was required to demonstrate statistically significant differences between the experimental group and the two control groups.

Chernikoff and Brogden (3) also investigated the effect of different instructions upon the occurrence of sensory pre-conditioning. For one set of subjects the following sentences were added to the instructions. *During the experiment the stimulus will change to a tone. You are not to respond to the tone.* For another set of subjects the added sentences were as follows: *During the experiment the stimulus will change to a tone. Do not decide in advance what you are going to do to the tone; do what seems natural.* The experimental groups given the above instructions did not give a frequency of response to the tone that was significantly different from the frequency of response to the tone made by their respective control groups. Two additional experiments on sensory pre-conditioning have been completed, the results for both of which were largely negative.[1] One was a two-factor experiment in which the number of pre-conditioning trials (2, 10, 15, 20) and the number of reaction time trials to the light (10, 20, 30, 40, 50) were varied. The experimental group that received the exact conditions of the experiment by Brogden (2), 10 pre-conditioning trials and 30 reaction time trials to the light, showed

* Supported in part by the Research Committee of the Graduate School from special funds voted by the State Legislature.

[1] Mr. Rube Chernikoff collaborated with the author in the first of these experiments and Mr. Lee Gregg collaborated in the second experiment.

the phenomenon by giving a response to the post-test trials of tone that was of significantly greater frequency than its control groups. The remaining experimental groups, with other quantities of these two variables, gave inconsistent results. There were no systematic quantitative changes in sensory pre-conditioning either as a function of the number of pre-conditioning trials or as a function of the number of training trials to the light. The second experiment dealt only with the number of pre-conditioning trials (2, 5, 10, 20); the number of training trials to light was held constant at 30. The experimental groups that received 5 and 10 pre-conditioning trials showed a greater frequency of response to the post-test trials of tone than did their respective control groups, but the differences were not statistically reliable. The other experimental groups gave a frequency of response in the post-test that was either below or between the frequencies shown by the control groups.

The inconclusive results of the above experiments might be made conclusive by increasing the number of *S*s; *N* for each group in the first experiment was three and in the second was 10. However, because two control groups are required for each experimental group for each value of the independent variable, the present technique of sensory pre-conditioning has a very low efficiency. Instead of increasing the number of *S*s or of attempting to test the effect of other variables on the magnitude of sensory pre-conditioning, time spent in the discovery of a more efficient procedure should be a wiser expenditure of research effort. Since Chernikoff and Brogden (3) found that an indefinite localization of the tone increased the efficiency of the sensory pre-conditioning technique used by Brogden (2), the hypothesis that change of the light stimulus from one of specific locus to one of increase in the general illumination of the experimental room will increase the efficiency of the sensory conditioning procedure appears fruitful. The verbal reports of *S*s given the sensory pre-conditioning treatment indicate that many hypotheses about the treatment occur, many of which prevent the occurrence of response to the post-test trials of the tone. It was decided, therefore, to decrease the intensity of both the light and tone and to make both of indefinite locus. The decision to use a tone of weak intensity suggested the possibility of using the measurement of an absolute threshold to the tone both as the training procedure (conditioning to the tone) and to provide the setting in which a response to the light might be measured.

The present paper is a report of an experiment, the design of which is based on the above hypotheses. When an *S* of the experimental group reported for an experiment on auditory acuity, he listened to 10 trials of a 1000-cycle tone approximately 15 db. above the threshold. Each time the tone was presented, the illumination of the room was increased slightly. *S* was then instructed to press a telegraph key as soon as he heard the tone. He was also told to signal *E* by pressing the key even when he was in doubt about hearing the tone. The experimental treatment consisted of presenting 30 trials of the tone at the same intensity level as that of the tone given during the sensory conditioning trials; then 10 trials of the light alone interspersed at random with 10 trials of the tone; and finally an absolute threshold was determined for the tone with the light in combination, and for the tone alone. With this procedure, only one control group is required, the *S*s of which are given identical treatment to that given *S*s of the experimental group except that the initial treatment of 10 trials of tone alone replaced the 10 trials of tone and light in combination.

Experimental Procedure

The experimental room and apparatus are essentially the same as that described by Cherni-

koff and Brogden (3). *S* sat at a table facing an X-ray mirror on the wall of the sound-insulated, light-shielded experimental room. The surface of the table contained only a telegraph key that was attached to the right of *S*. A pair of two-in. loud speakers was fastened beneath the table, one on each side. Since the right and left speakers were an equal distance from *S*s head, there was an indeterminate localization of the source of the 1000-cycle tone stimulus. A 20-watt fluorescent light fixture was attached to the ceiling directly behind the standard light fixture and directly behind *S*. The light fixture was activated at a low voltage level prior to the experiment and thus provided a part of the general illumination of the room throughout the experiment. The light stimulus was an increase in the general illumination of the room produced by an increase in the voltage across the fluorescent tube.

The rest of the apparatus was located in the control room, from which point the stimuli were presented and the occurrence and latency of the key pressing signal was recorded. An electronic timer controlled the relays that presented the stimuli for a duration of two sec. The timer also started an electronic chronoscope that was stopped by closure of the telegraph key. The latency of response was measured in milliseconds. The 1000-cycle tone was produced by a Jackson audio oscillator (Model No. 652) and was fed to the loud speakers through a General Radio attenuation box (Type 249-H). The intensity of the tone at the locus of *S*'s head during the sensory conditioning trials and the training trials was 28 db. above 10^{-16} watts/sq. cm. and was measured by a General Radio Sound Level Meter, Type 759-B, with the weighting control set at the C position. The intensity of noise from the ventilation system of the experimental room was 60 db. above the reference level at the locus of *S*'s head. The 1000-cycle tone was always heard against the background of noise. The illumination of the room was measured at the locus of *S*'s head as the brightness of the reflection from the mirror. The normal brightness was measured by means of a Macbeth Illuminometer at 2.66 millilamberts. When the light stimulus was on, that is when the voltage across the fluorescent tube was increased, the brightness of the reflection from the mirror at the locus of *S*'s head was 3.40 millilamberts. The difference in brightness for these two conditions of 0.74 millilamberts is therefore the brightness of the light stimulus.

The *S*s of the experiment were volunteers from the classes in elementary psychology. Forty *S*s, 28 men and 12 women, were assigned at random to the experimental and control groups, so that *N* for each group was 20. When *S* kept his appointment for an experiment on auditory acuity, he was seated at the table of the experimental room, and asked to listen carefully to a series of weak tones. If his assignment was to the experimental group, 10 trials of the tone and light were given. If assignment was to the control group, 10 trials of the tone alone were presented. If *S* did not indicate to *E* that he heard the tone on the first trial, he was asked if he heard it. If *S* did not hear the tone, *E* increased the intensity by 10 db. Only a few *S*s needed this treatment, and all were able to hear the tone of standard intensity (28 db. above 10^{-16} watts/sq. cm.) after two trials of the more intense tone. When the 10 trials of listening had been completed, *E* instructed *S* to signal by pressing the telegraph key as soon as the tone was heard. *S* was further instructed that some tones would be too weak to be heard and that others would be heard only part of the time; and that he should signal by pressing the key, even though he might be doubtful about hearing the tone. *E* then went to the control room and presented 30 trials of the tone, then 10 trials of the light alone and 10 trials of the tone in a random order. Following this, two measurements of the absolute threshold were made, one with the light occurring each time a tone was presented and the other without the light. The odd-numbered subjects of the experimental and control groups each make up a subgroup for which the threshold with the light was taken first. The even-numbered subjects are in the subgroups for which the threshold with the light was taken last. The threshold was measured in each case by a modified form of the method of constant stimuli and the method of limits (4). Five decibel steps between stimulus intensities were used. The test values were selected so that 100 percent response was obtained for the loudest and zero percent for the weakest tone. Four steps were usually sufficient to obtain the threshold, the intensity of tone for 50 percent response. All thresholds are expressed in db. difference from the standard reference level of 10^{-16} watts/sq. cm.

Results

The design of the present experiment is such that two possibilities for measuring the effect of sensory conditioning exist. One is a difference between the experimental and control groups in terms of response to the trials of the light alone. The other is a difference in auditory acuity between the light and no-light condi-

TABLE I

The Results from the Two Methods of Measuring Sensory Conditioning

The odd-numbered *Ss* make up the sub-groups for whom the auditory threshold with light was made first and the even-numbered *Ss* are the sub-groups for whom the auditory threshold with light was second. The thresholds are expressed in db. above the reference intensity of 10^{-16} watts/sq. cm. The standard errors of the sub-group and group mean thresholds are 0.63 db. and 0.45 db., respectively. Both are derived from the error mean square of the analysis of variance presented in Table III.

Subject Number	Experimental Group			Control Group		
	Frequency of Response to Light	Threshold with Light	Threshold without Light	Frequency of Response to Light	Threshold with Light	Threshold without Light
1	0	18.0	20.5	1	20.5	20.5
3	0	20.5	18.0	0	25.5	25.5
5	0	8.0	13.0	0	15.5	20.5
7	0	20.5	23.0	0	15.5	18.0
9	0	18.0	20.5	0	25.5	20.5
11	0	13.0	18.0	0	15.5	15.5
13	0	20.5	23.0	0	18.0	13.0
15	0	27.5	27.5	0	13.0	10.5
17	0	−4.5	5.5	0	15.5	15.5
19	0	15.0	15.0	0	20.5	20.5
Sub-group Mean	0.00	16.50	18.75	0.10	18.50	18.00
2	0	20.5	25.5	0	20.5	18.0
4	0	25.5	25.5	0	20.5	20.5
6	0	15.5	20.5	0	25.5	25.5
8	0	13.0	15.5	0	18.0	23.0
10	0	15.5	18.0	0	20.5	20.5
12	1	15.5	18.0	2	15.5	18.0
14	2	18.0	15.5	0	13.0	13.0
16	0	20.5	23.0	0	18.0	20.5
18	0	10.5	10.5	1	23.0	25.5
20	0	10.5	15.5	0	30.5	30.5
Sub-group Mean	0.30	15.50	18.25	0.30	20.50	21.50
Group Mean	0.15	16.00	18.50	0.20	19.50	19.75

tions of threshold measurement. Table I presents these basic data for each *S* in the experimental and control groups. An examination of Table I shows that only five *Ss* responded at all to the light, that two of these were members of the experimental group and three were members of the control group, and that none who responded did so with a frequency greater than two. This particular measure provides no evidence whatever of any sensory conditioning.

The threshold, data however, appear to show positive results. The mean threshold with light is 16.00 db. and without light is 18.50 db. for the experimental group. The mean thresholds for the control group are 19.50 db. with light and 19.75 db. without light. The reliability of the increase in auditory acuity produced by the light in *Ss* of the experimental group was tested by analysis of variance. The experimental design for obtaining thresholds with and without light is

that of a 2 × 2 latin square both in the case of the experimental group and the control group. The diagonals of the square represent the treatment, presence of light with the tone versus no light, the columns of the square represent the ordinal position of treatment, and the rows represent the sequence of treatment. The procedure for analysis of variance is that devised by Grant (5) and provides a pure error estimate from which the effects of the diagonals, columns, rows, and individual variation within each cell of the latin square have been removed. The results of these analyses are presented in Table II. Row 6 of Table II shows that the effect of light versus no light in obtaining an auditory threshold is statistically significant at better than the five percent level of confidence in the case of the experimental group, but is not significant for the control group. The ordinal position of the treatment has no significant effect on the auditory threshold for either group (row 7). The sequence of treatment has a significant effect for the control group but not for the experimental group (row 8). However, since the *F*-ratio of the mean square for sequence of treatment to the mean square for individual variation is significant (row 9), the conclusion may be drawn that the significant effect found for sequence of treatment may be accounted for in terms of individual variation alone.

Since neither the sequence nor ordinal position of the light-no-light condi-

TABLE II

RESULTS FROM ANALYSIS OF VARIANCE OF THE 2 × 2 LATIN SQUARES OF THE AUDITORY THRESHOLD DATA FOR THE EXPERIMENTAL AND CONTROL GROUPS

Source	*d.f.*	Experimental Group		Control Group	
		Sums of Squares	Mean Square	Sums of Squares	Mean Square
1. Treatment	1	62.50	62.50	0.63	0.63
2. Ordinal Position of Treatment	1	0.63	0.63	5.62	5.62
3. Sequence of Treatment	1	5.63	5.63	75.62	75.62
4. Individual Variation of *Ss* Within Sequence	18	1290.62	71.70	725.00	40.28
5. Error	18	80.62	4.48	62.50	3.47
F-Ratio		*F*	*F* at 5% level of confidence	*F*	*F* at 5% level of confidence
6. $\frac{\text{Mean Square Row 1}}{\text{Mean Square Row 5}}$		13.95	4.41	0.02	4.41
7. $\frac{\text{Mean Square Row 2}}{\text{Mean Square Row 5}}$		0.01	4.41	1.62	4.41
8. $\frac{\text{Mean Square Row 3}}{\text{Mean Square Row 5}}$		1.26	4.41	21.79	4.41
9. $\frac{\text{Mean Square Row 4}}{\text{Mean Square Row 5}}$		16.00	2.19	11.61	2.19
10. $\frac{\text{Mean Square Row 3}}{\text{Mean Square Row 4}}$		0.07	4.41	1.88	4.41

TABLE III

RESULTS FROM ANALYSIS OF VARIANCE OF THE OVERALL AUDITORY THRESHOLD DATA

Source	d.f.	Sum of Squares	Mean Square	F	F at 5% level of confidence
Sensory conditioning treatment	1	112.82	112.82	28.35	4.12
Threshold treatment	1	37.82	37.82	9.50	4.12
Interaction	1	25.32	25.32	6.36	4.12
Individual variation of Ss within sub-groups and within groups	36		55.99	14.07	1.76
Error	36		3.98		

tion was found to affect the thresholds of either the experimental or control groups, an overall analysis of variance of the two-factor experimental design, sensory conditioning training versus no sensory conditioning training and thresholds obtained with light versus those without light, may be made that uses the previous analyses.[2] The results of this analysis are presented in Table III. The error mean square is the mean of the error mean squares obtained from the analyses of the 2 × 2 latin squares and is a pure error estimate from which the effects of all other variables shown in Table III have been eliminated. The mean square for individual variation of *Ss* within sub-groups and within groups is the mean of the mean squares for individual variation of *Ss* within sequence from the analyses of the 2 × 2 latin squares. The other mean squares were obtained by the standard procedure (7). All variables produce an effect significant at better than the five percent level of confidence. However, since the *F*-ratio of the mean square for individual variation of *Ss* within sub-groups and within groups to the mean square for sensory conditioning (2.02) is not significant (*F* at five percent level of confidence = 250.09), the significant effect for sensory conditioning treatment, that is, the difference between the mean for the thresholds both with and without light for the experimental group (17.25 db.) and the mean for the thresholds both with and without light for the control group (19.63 db.), may be attributed to individual variation alone. The significant effect for threshold treatment, the difference between the mean threshold with light for *Ss* of both groups (17.75 db.) and the mean threshold without light for *Ss* of both groups (19.13 db.), is probably a sole function of the effect of the sensory conditioning treatment on the threshold of *Ss* in the experimental group. However, the significance of the interaction term provides the critical measure of positive results by the experiment. This demonstrates a reliable difference (−2.25 db.) between the mean difference in threshold with light and threshold without light for the experimental group (−2.50 db.) and the mean difference in threshold with light and threshold without light for the control group (−0.25 db.).

DISCUSSION

In terms of the standard procedure for measuring sensory pre-conditioning, the results of the present experi-

[2] This procedure for analysis of variance was divised by Dr. David A. Grant, to whom the author expresses grateful appreciation.

ment are entirely negative. Positive results were obtained, however, by means of a new procedure. The presence of a light with tones presented in determination of an auditory threshold increases sensitivity over that obtained by threshold determination with the tone alone, when the tone and light have been presented together at a prior time. The difference in results on auditory acuity with and without the presence of the light for the experimental and control groups makes it clear that the initial trials of the tone and light in combination are responsible for the greater auditory acuity in the presence of light shown by *Ss* of the experimental group. This is a direct statement of the results of the experiment. Sensory conditioning between the tone and light in the present experiment has status only as an inference, as is the case also for the experiments on sensory pre-conditioning (1, 2, 3, 6). The inference is drawn from the effect that the trials of tone and light in combination have upon auditory acuity when the light is presented simultaneously with the tone. The term sensory conditioning refers to a relation between the tone and light that is a function of the repetitive pairing of the two. Since the nature of this relation can only be speculated upon, the phenomenon demonstrated by an increase in auditory acuity and the phenomenon demonstrated by response to the stimulus for which specific training was not given, may or may not represent the same relation between the tone and the light. Although in both the present experiment and the earlier experiments (2, 3), there were ten sensory conditioning trials, the context produced by the instructions was markedly different. Also, in the present experiment, the light stimulus was general as was the tone, and both were of lesser intensity than was the case for the earlier experiments. The lack of response to the trials of light alone in the present experiment might serve as an argument against the relation between the tone and light being the same as that which occurred in the experiments where its sole measurement was in terms of frequency of response to the other stimulus.

The possibility also exists in the present experiment that the phenomenon is of the kind usually spoken of as intersensory relations. The latter term would be most applicable if the presence of the light with the tone served to increase auditory acuity independently of the initial treatment of combining the tone and light. Because this treatment is essential to the effect of the light on auditory acuity, the phenomenon is taken as an indicator of sensory conditioning. Of course, the relations between stimuli that are involved in intersensory relations may be of the same nature as that involved in sensory conditioning.

Summary and Conclusions

A new procedure for measuring sensory conditioning between a tone and light was tested. *Ss* of the experimental group reported for an experiment on auditory acuity and were asked to listen to a weak tone that was accompanied by a small increase in the general illumination of the room. *Ss* of the control group were given the same tone without the light. Both groups were instructed to signal *E* that the tone was heard by pressing a telegraph key. The *Ss* were then given 30 trials of the tone and then 10 trials of the light and 10 trials of the tone in random order.

Finally, one auditory threshold was obtained when the light was present with the tones, and another threshold was obtained without the light.

Few *Ss* responded with key pressing to the trials of light and there was no difference between the experimental and control groups in this respect. However, the *Ss* of the experimental group showed a significant increase in auditory acuity when the light was present over their auditory acuity without the light. The control group showed no difference in auditory acuity. Since the gain in auditory acuity by the experimental group is a function of the initial trials of the tone and light together, the inference is drawn that some relation between the two, a sensory conditioning, occurred. The phenomenon of the present experiment is discussed with reference to the phenomenon of earlier studies on sensory pre-conditioning.

(Manuscript received June 27, 1949)

References

1. Brogden, W. J. Sensory pre-conditioning. *J. exp. Psychol.*, 1939, **25**, 323–332.
2. Brogden, W. J. Sensory pre-conditioning of human subjects. *J. exp. Psychol.*, 1947, **37**, 527–539.
3. Chernikoff, R., & Brogden, W. J. The effect of instructions upon sensory pre-conditioning of human subjects. *J. exp. Psychol.*, 1949, **39**, 200–207.
4. Culler, E., Finch, G., Girden, E., & Brogden, W. Measurements of acuity by the conditioned-response technique. *J. gen. Psychol.*, 1935, **12**, 223–227.
5. Grant, D. A. The statistical analysis of a frequent experimental design. *Amer. J. Psychol.*, 1949, **62**, 119–122.
6. Karn, H. W. Sensory pre-conditioning and incidental learning in human subjects. *J. exp. Psychol.*, 1947, **37**, 540–544.
7. Snedecor, G. W. *Statistical methods.* (3rd ed.) Ames, Iowa: Collegiate Press, 1940.

THE RELATION OF VERNIER AND DEPTH DISCRIMINATION TO WIDTH OF TEST ROD[1]

BY RICHARD N. BERRY

Indiana University

AND

LORRIN A. RIGGS AND WALTER RICHARDS

Brown University

It has been previously established that the angular *height* of test objects is a determinant in depth discrimination and in vernier discrimination (1, 2, 4). The present experiment was designed to evaluate the effects of varying the angular *width* of test objects in both a depth discrimination situation and a vernier discrimination situation.

Apparatus and Procedure

The apparatus and procedure employed here have been described in detail elsewhere (3). The essential parts of the apparatus were two vertical blackened steel rods placed one above the other in an otherwise homogeneous white field. The desired variation in thickness of line was achieved by the use of rods of various diameters. These vertical rods were placed 4622 mm. in front of the subject's eyes in the median plane. In the case where a vernier or lateral displacement was to be judged, the lower rod was moved laterally by the use of a micrometer adjustment. In the case where a real depth judgment was to be made the lower rod was moved in the median plane. In each situation the subject was instructed to judge the lower rod in relation to the upper. In the vernier situation a response of 'Right' or 'Left' was required, and in the real depth situation 'Front' or 'Back' was the required response. A constant stimulus method was used, each threshold[2] being computed on the basis of 100 responses. In both the vernier and the real depth situations, two such thresholds (for each subject) were determined for each of five different diameters or widths of the test rods. These five diameters (in sec. of visual arc) were: 26.8, 53.5, 107, 214, and 424. (In mm. the same diameters were 0.6, 1.2, 2.4, 4.8, and 8.5.) In all cases the vertical distance between the upper and lower rods was 134″ of arc

TABLE I

The Two Thresholds for Each *S* for Five Widths of the Test Objects in the Vernier and in the Real Depth Situations

Angular Width (in secs.)	Vernier			
	L. A. R.	R. N. B.	W. R.	Average
26.8	3.87 2.60	2.49 3.31	3.40 2.79	3.08
53.5	2.68 3.06	3.31 3.59	3.74 2.40	3.13
107	3.29 2.14	2.73 2.14	4.05 2.51	2.81
214	2.88 2.57	2.31 2.46	3.01 2.82	2.68
424	3.81 3.24	2.79 4.26	3.33 2.75	3.37
	Real Depth			
	L. A. R.	R. N. B.	W. R.	Average
26.8	4.60 6.08	2.60 2.29	2.95 5.95	4.08
53.5	7.99 10.20	2.34 2.42	3.78 4.47	5.20
107	5.95 3.08	1.86 1.94	5.34 6.86	4.17
214	4.86 2.73	2.21 2.38	4.30 3.82	3.39
424	3.21 7.94	1.39 1.60	4.56 5.16	3.98

[1] These experiments were performed in the Psychological Laboratory of Brown University.

[2] The computation of thresholds was traditional and was the same as the one used previously in this situation (3).

(or 3 mm.), and the apertures in the apparatus were so arranged that the distance from the bottom of the lower rod to the top of the upper one subtended a visual angle of 1° 30′.

The subjects were all male adults, two of whom had had considerable previous practice in the general experimental situation. The third subject (W. R.) had not had such previous practice. The results do not vary in such a way that previous practice can be construed as a significant variable. All of the three subjects had normal, uncorrected vision, with the exception of an exophoria in the case of L. A. R.

Results and Discussion

The results of the experiment are summarized in Table I and Fig. 1. There were only six thresholds (two for each of the three subjects) used in the determination of each point in Fig. 1 (taken from the last column of Table I). It might be expected that the differences between successive points (means) on either curve A or curve B would not be significant at the five percent level of confidence. An analysis of variance shows this expectation to be realized.[3] Since this is the case we may consider both lines as being essentially linear with a slope of zero.

Thus we may conclude that a change in width of the test object had no significant effect on either vernier or depth discrimination.

It is perhaps not surprising that the length of line should be an important determiner of vernier and depth thresholds (1, 2) whereas their width is shown to be relatively unimportant. For regardless of width, the subject can always judge vernier or depth alignment by reference to either or both of the straight edges of the test objects. Hecht (6) and Fry (5) among others have demonstrated the predominant influence of the edge of

[3] The analysis of variance did show a significant difference at the one percent level of confidence *between subjects* in the real depth situation, but not even at the five percent level in the vernier situation.

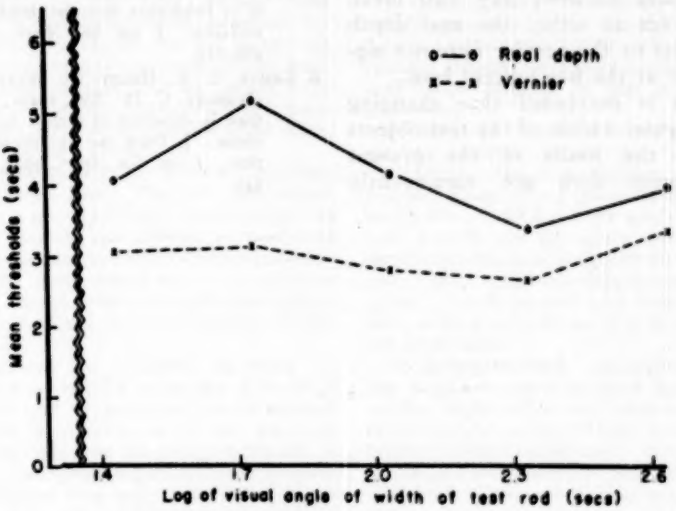

Fig. 1. Mean thresholds of three subjects for different widths of the test rods in the real depth and in the vernier situations. Each point is a mean of six values, since each subject's threshold was determined twice for every condition.

the contour rather than the area in determining acuity.

Summary

The present experiment was designed to test the effect of varying the width of the test objects in the vernier and real depth situations. Three adult male subjects viewed two vertical black test rods, 4622 mm. from the eyes, against a homogeneous white background. The lower rod was moved either in the frontal or in the median plane and was judged in reference to the upper rod. Five widths of test rods were used: 26.8, 53.5, 107, 214, 424 sec. of visual arc. The total angular height of the test object was constant at 1° 30′.

1. The means of the subject's thresholds varied from 2.68″ to 3.37″ in the vernier situation.

2. The means of the same subject's thresholds varied from 3.39″ to 5.20″ in the real depth situation. The differences between any two mean thresholds in either the real depth situation or the vernier were not significant at the five percent level.

3. It is concluded that changing the angular widths of the test objects within the limits of the present experiment does not significantly affect either vernier or the real depth thresholds.

4. It is inferred that the alignment of two rods may be accomplished by the use of either or both of the straight edges of each rod, regardless of the horizontal separation of these edges (i.e., width of the rod).

(Manuscript received June 15, 1949)

References

1. Anderson, E. E., & Weymouth, F. W. Visual perception and the retinal mosaic: I. Retinal mean local sign: an explanation of the fineness of binocular perception of distance. *Amer. J. Physiol.*, 1923, 64, 561–594.
2. Averill, H. L., & Weymouth, F. W. Visual perception and the retinal mosaic: II. The influence of eye movements on the displacement threshold. *J. comp. Psychol.*, 1925, 5, 147–176.
3. Berry, R. N. Quantitative relations among vernier, real depth and stereoscopic depth acuities. *J. exp. Psychol.*, 1948, 38, 708–721.
4. French, J. N. The unaided eye: Part III. *Trans. opt. Soc.*, 1919/1920, 21, 127–156.
5. Fry, G. A. The relation of the configuration of a brightness contrast border to its visibility. *J. opt. Soc. Amer.*, 1947, 37, 166–175.
6. Lamar, E. S., Hecht, S., Shlaer, S., & Hendley, C. D. Size, shape, and contrast in detection of targets by daylight vision. I. Data and analytical description. *J. opt. Soc. Amer.*, 1947, 37, 531–545.

AN EXPERIMENT ON THE EXPRESSIVENESS OF SHELL AND TEXTILE 'MONTAGES'

BY IVY G. CAMPBELL-FISHER

Department of Aesthetics (Psychological and Philosophical), Wells College *

Will observers agree as to what is expressed by the sense materials of art (color, line, tone, etc.) and their compositions? It would have been argued by the experimenter that large agreement indicates that judgments were based on something objectively present in the stimulus (its intrinsic expressiveness), while disagreement might tend to support the view that judgments were made by the observers on some subjective basis, e.g., their varied associations or conditionings.

In attempting an experimental answer to the question, particularly an answer directly applicable to aesthetics, it is difficult to approximate the conditions of art and yet avoid the preconceptions apt to be connected with an actual work of art. The experimenter therefore set up for observation, on nine small separate tables covered with papers or textiles, compositions of sea shells and other sea jetsom. In each the attempt was made to express one certain mood. While these 'montages' did not pretend to be works of art, they bore certain relations to them, especially that of being composed of the colors, shapes, lines, etc., of visual art.

The montages were set up in two rooms of the experimenter's house. Observers were invited to come in at various times. Artificial light was kept on to make this condition as constant as possible whether the observations were made day or night. Some description of the montages follows, together with the experimenter's mood-title for them and the letter each was marked with and by which the observers referred to them:

A. *Aggressiveness.* On burnt orange silk were placed an upright starfish, fish vertebrae, shells with sharp points, various sharp-angled cutouts of mat black paper.

C. *Aloofness, with a touch of pride.* Above silver paper, a polished chambered nautilus lay on a black velvet cutout which somewhat resembled an arrowhead pointed away from the observer. Beside this were three flat, round clam shells, one upright, about three and a half inches in diameter, with the pearly inner side exposed.

D. *Mood of jitterbug-ish swing.* On bright red-orange Chinese paper with small irregular silver patches, across which lay an angular-embroidered Tepic Indian belt in black and brown, were scattered many sharp shells covered with jagged lines. Three fish skulls and a spray of maroon sea weed stood upright in back against a bit of straw weaving with sharp-angular and square designs in black, white, orange, and tans.

E. *Gayety (Gentle).* On dullish yellow wool and silk were placed groups of yellow buttercup shells and red and yellow small scallops before a mounting background of large translucent pen-shells (point upreared) in pale yellow with bluish tints.

F. *Depression.* On black velvet were placed one long recumbent branch of a dark brown bushy sponge, other brown driftage, one much weathered conch with a remainder of dull pink inside, and a black and white conch fragment.

G. *Serenity.* On pale green sequins lay a few whitish, flat, lined shells and one small pear shell, faintly tinged off-white toward yellow-brown. Upright at each back corner stood a spray of sea grape with two or three large dried pale green leaves.

H. *Joyful exuberance.* On bright gold Chinese paper were placed groups of small spotted cockles, bright yellow and white snail shells, brilliant red lion paws, and various other bright (mainly yellowish) small shells. Behind, on the wall, hung a soft East-Indian silk scarf of yellow with bright variegated designs, particularly in red.

I. *Apprehensive indicision.* A white linen band lay across black silk, and over both were placed black, white, or gray shells, some forming

* This was the experiment on which Dr. Campbell was working at the time of her death in April, 1948. It is reported by her husband from her conversation, notes, and the data. Since the 'montages' are hardly duplicable, no detailed description of them is offered; however, color films of them were made, and if anyone wishes more exact information these will be loaned him (or further particulars furnished) by Lewis Fisher, Childwold, New York, or Captiva, Florida.

a line of broken angles. A spray of dried white-to-purple seaweed stood at one side in the rear.

J. *Constancy.* On dull blue paper (approximately Munsell PB 4/4), with lighter dull blue paper at either side, were placed four white chalk-textured shells and five butterfly-shaped double shells exposing a white and blue opalescent texture.

(Originally there had been a montage B, *Bourgeois complacency.* But since this seemed more a concept than a mood, and, perhaps for this reason, tended to be apparent to adults but to confuse younger observers, it was eliminated before the experiment proper began.)

The above order is that in which an observer, if he began at a certain point and ended at another (which he was not told to do), would encounter the montages. Observers were given a sheet with the above names in the order E, C, I, A, F, G, H, J, D (as in Table I). Each name was accompanied by a blank on which the observer was to enter the letter that designated the montage he choose for it. The observers were handed the following typed instructions:

Montages In Search of Their Titles

I have hoped that each of these montages expresses some over-all mood. Supposing that the words given you on a separate paper name these moods, choose for each of the montages that name which seems to you to fit it best. *No name should be assigned to more than one montage.* Make it your task to assign a different name to *every* one. Try to choose on the basis of an over-all mood, not on the basis of any association with parts. Indicate your choice by *entering the letter* that is on the floor near each montage by the name you think best fits that one. Look at all before you make any choice.

The observers walked freely about and spent whatever time they chose, generally from 10 min. to a half hour, in selecting the montages to go with the various titles. When they were through, most of them wanted to know the experimenter's choice of titles. The subject was freely discussed with them, on the condition of secrecy.

The observers consisted of 41 Wells College students (women), six 'boy-friends,' and a miscellaneous group fo 21 non-students called here 'adults,' who were all older persons except that one seventh-grade girl presented herself and is included in this group. The total was 68.

Table I shows the title given by the experimenter to each montage, and the number of observers who gave it that and the various other titles. Of the 68 observers, the largest agreement is that of 63 on *Depression.* With five of the nine montages, half or more of the observers gave the title selected by the experimenter. In only two of the nine cases did the experimenter's choice of title fail to receive more than twice as many votes as any other one title. In only one case was some other mood name chosen by more observers than chose the experimenter's title. (This was *Aggressiveness,* which 18 called that

TABLE I

NUMBER OF SUBJECTS ASSIGNING EACH TITLE TO EACH MONTAGE

In the left-hand column appear *E*'s titles for each montage together with the letter by which *O* referred to it. The other vertical columns show the titles the 68 *O*s selected for each montage. The numbers italicized indicate agreements between *E* and *O*.

E's title for montages	Gayety	Aloof	Ap. Ind.	Aggres.	Depr.	Ser.	Joy	Const.	Jit.
E. Gayety (Gentle)	*37*	1	3	2	0	2	17	4	2
C. Aloofness, with a touch of pride	3	*37*	6	3	1	13	0	5	0
I. Apprehensive indecision	1	0	*31*	14	0	10	1	2	9
A. Aggressiveness	1	11	11	*18*	3	3	1	1	19
F. Depression	0	2	1	1	*63*	0	0	1	0
G. Serenity	6	12	3	0	1	*34*	3	9	0
H. Joyful exuberance	14	0	4	5	0	0	*32*	1	12
J. Constancy	5	5	5	0	0	6	2	*45*	0
D. Mood of jitterbug-ish swing	1	0	4	25	0	0	12	0	*26*

TABLE II

EXTENT OF AGREEMENT BY SUBJECTS WITH THE EXPERIMENTER'S TITLES

The score refers to the number of titles chosen by *O* in agreement with *E*'s titles. A score of 9, e.g., means complete agreement. Under the number and classification of *O*s appears the number of *O*s making the scores indicated in the right-hand column. The number 8 was an impossible score, and no *O* was in complete disagreement with *E*. The abbreviation 'c.c.' = coefficient of contingency.

Score	c.c.	19 Underclassmen	22 Upperclassmen	41 Both	6 'Boy Friends'	21 Adults	68 Total
9	.943	3	2	5	0	1	6
8	—	—	—	—	—	—	—
7	.904	3	5	8	1	2	11
6	.870	2	0	2	0	2	4
5	.817	4	5	9	1	2	12
4	.727	4	5	9	2	3	14
3	.577	3	4	7	1	4	12
2	.333	0	1	1	1	5	7
1	.000+	0	0	0	0	2	2
0	.000	0	0	0	0	0	0
Median c.c.		.817	.817	.817	.727	.652	.727
Average c.c.		.799	.762	.779	.681	.585	.711

while 19 called it *Mood of jitterbug-ish swing*. It was between these somewhat similar montages that the major disagreement appeared; for instance, among observers making only two 'errors,' the 'confusion' between *Aggressiveness* and *Jitterbug* occurred five times while six other errors occurred only once each. The experimenter could not claim to be an authority on jitterbug—which in itself, moreover, may be quite related to aggressiveness.)

Table II shows the scores and coefficients of contingency for groups into which the observers may be divided. The median c.c. for the entire group was .727, the average c.c. was .711. It was possible for the observers to have nine titles in agreement with the experimenter, impossible to have eight, and none achieved zero. It will be noted that the underclassmen had three 'perfect scores,' the upperclassmen two, and the adults one. The median c.c.'s for the three groups are respectively .817, .817, and .652. The average c.c.'s are .799, .762, and .585.

While, within the limits of the number of observers, the c.c. for all and the c.c.'s for each group show significant agreement with the expression of the montages intended by the experimenter, the progressively lesser agreement among underclassmen, upperclassmen, and adults suggests the possibility that with increasing age, education, and the gathering of associations, our sensitiveness to an originally intrinsic expressiveness of sense materials becomes less accessible. Among the adult observers, some of the most art-experienced made the 'poorest' score, while the one 'perfect' score in this group was that of an observer of comparatively little familiarity with art.

In considering this possibility, however, it should be noted that between the two groups of student observers there was a difference in addition to their being upper- and underclassmen. All of the experimenter's students in aesthetics and a number of other students specifically interested in art had been used as preliminary observers and helpers in arriving at the final

arrangements of and titles for the montages. Since these students could not be used as observers in the actual experiment, and since they were all upperclassmen, a certain number who may be thought to be the more sensitive were automatically eliminated from that group. The underclassmen observers underwent no such combing-out process.

Conclusions

A group of shell and textile montages, in which so far as associations having to do with being cast up by the sea were concerned these associations were shared by all of them, was with significant agreement judged by 68 observers to express each its own mood. This looks toward an *objective* or *intrinsic expressiveness* of the sense materials of art and their compositions,[1] and tends to cast doubt on the widely held belief that art's expressiveness depends subjectively upon an individual's personal reactions, though these personal reactions may well be strong enough to lead to differences of judgment in individual cases.

There is also some indication that the more naive observers may be closer to the direct expressiveness of a composition, that the ability to recognize this may decrease under a gathering overlay of experience.

(Manuscript received July 18, 1949)

[1] Dr. Campbell had long been collecting data for a work on aesthetics, sections of which would have been entitled *The Logic of Sense*, and *Intrinsic Expressiveness*. Cf. forthcoming posthumous publications in *The Journal of General Psychology*.

AN HYPOTHESIS CONCERNING THE GENERATION AND USE OF SYNONYMS [1]

BY WILLIAM M. LEPLEY

The Pennsylvania State College

Introduction

The experiments to be reported were planned and executed as tests of an hypothesis that the principle of reactive inhibition, as employed by Hull (2) and others, operates in the generation and use of synonyms. This principle is simply expressed in Hilgard's (1) modification of Hull's Postulate 8: "The evocation of any reaction generates reactive inhibition. Reactive inhibition is spontaneously dissipated in time." [2]

By way of restatement and amplification, this principle implies that when any reaction occurs, its reaction threshold is raised and the probability of its later occurrence is reduced. It would supposedly follow that the probability of later occurrence would be reduced progressively with continued repetition and the rising of the reaction threshold. It is proposed that this rise in a particular response threshold would leave the reaction thresholds for other reactions, in the repertoire, relatively low and would enhance the probability of the occurrence of these reactions. An extension of this line of thought is here proposed; that repetitive reacting may ultimately raise the reaction threshold to the point whereat novel reactions, serving the same or similar ends, will appear and enrich the variety of the behavior repertoire.

Now, as a special case, involving the principles suggested above, it is proposed that when a concept is used repetitiously by a person, or the persons of a culture, the synonym vocabulary representing the concept will be enriched, and that this enrichment will be proportional to the repetitiousness or frequency of use.

Two studies appear to bear upon the present problem. Zipf (8) made a study of the meaning-frequency relationship for words in which he used Thorndike's word book (4) as his source for frequency-of-use, and Thorndike's dictionary (5) as his source for number of meanings. His study clearly shows that, by these criteria, the more frequently used words have the larger numbers of meanings.[3] In a similar study, Thorndike (7), using the later edition of his word book (6) as his source for frequency-of-use, and The Shorter Oxford English Dictionary (3) as his source for number of meanings and for date of entrance to the language, showed that: 1. There is a relationship between the date of a word's entrance to the language and its number of meanings for words of equal frequency of occurrence, the older words having the greater number of meanings. 2. There is a relationship between frequency of occurrence and number of meanings for words of equal length of time in the

[1] This writing is an elaboration of a paper, of the same title, read before the 1949 Convention of the American Psychological Association.

[2] It will be noted that the point concerning work (*W*) in Hull's Postulate 8 is omitted in this simplification. It is judged that the variable of work, or effortfulness-of-task, is irrelevant to the present study.

[3] Zipf, in his recent book (9), submits other data as evidences of language changes, all of which he interprets in terms of the principle of least effort.

language, the more frequently occurring words having the greater number of meanings. These two studies are here interpreted as exhibiting phenomena, in a sense 'obverse' to the phenomena predicted and sought in the present study. In the studies cited, the word-form may be thought of as remaining constant, while the conceptual aspects vary with frequency of use. (It seems reasonable to suppose that the cumulative frequency of usage of a word varies directly as the length of time the word has been in the language.) Whereas in the present study the *concept* is thought of as remaining constant and the word-forms (synonym vocabularies) are expected to vary.

The Experiments

Five replications of a simple experiment were executed as tests of the hypothesis under consideration.

Subjects.—Five undergraduate classes, with *N*'s of 39, 51, 19, 22, and 44, were used. Pains were taken that all subjects remained appropriately naive until all five classes had been dealt with.

Procedure.—The following instructions were read to each class:

> Write these four words, evenly spaced, across the top of the sheet of paper provided. Do not put your names on your papers. [In Experiment I the four words were: HORSE, DOG, CHICKEN, and SHEEP. A different set of words was used for each replication as is shown in Table I.] The order in which you write them is not important. [Ample time was given for writing.] Now, I want you to assign a rank order number to each word; this based upon your estimate of the relative frequency with which you use these words. By *use* we mean *speaking*, *writing*, and *thinking*. The word you use *most* frequently in *speaking*, *writing* and *thinking* should be given the rank of one. The word you use *least* frequently should be given the rank of four. Assign ranks to the other two words according to the same principle. [Ample time was allowed for the ranking procedure. It may be said that no subject appeared to have any difficulty in assigning ranks.] Now, under each word I want you to write as many synonyms for the word as you can. You may have as much time as you need. Synonyms are different words that have the same meaning. For example, synonyms for the word *baby* are: babe, infant, etc. For your synonymns you may write slang words, obscene words or words of any kind, but be careful that, as you use them, they actually have the same meaning as the word given. Work independently. Be certain that you do not see the words being written by those near you. Go ahead now and write as many synonyms as you can for each of the four words. [Sufficient time was allowed for the synonym writing, such that, upon questioning, no one requested more time. The time required for this phase varied from group to group, with an average of approximately seven min.] Now, turn your papers face downward and pass them to the ends of the rows for collection.

Two items of further explanation of procedures are in order. First, the five, four-word groups were obtained by soliciting graduate students and requesting:

> Give me four related words; four nouns, four verbs, four adjectives or four adverbs.

The first five word-groups conforming to these specifications were used. Second, the question might be raised as to why the instructions to the subjects were couched in terms of *words*, rather than in terms of *concepts*. The answer is that it is relatively more difficult to define the term *concept* for the understanding of undergraduate students.

The procedures used in all replications were identical except that each of the five groups used a different set of words.

Treatment of the Data and Results

1. There was no editing of the synonym responses, though many rather odd and, to the investigator, meaningless responses appeared in the records. Each synonym-response was counted.

2. For each respondent, the experimenter assigned rank order numbers to the words, these based upon the number of synonymous expressions produced. The word with the largest number of synonyms was given the rank of one.

3. For each respondent a rank-difference coefficient of correlation[4] was computed between the rank order based upon the respondent's estimated frequency of use and the rank order as

[4] The conventional, Spearman, rank-difference formula was used.

TABLE I

TABLE OF RANK FREQUENCIES
Ranks are Based on Estimated Frequency of Use

Experiment	Word	Ranks				N
		1	2	3	4	
I	Horse	2	16	17	4	39
	Dog	37	0	2	0	
	Chicken	0	20	13	6	
	Sheep	0	3	7	29	
II	Bored	1	9	18	23	51
	Tired	32	16	2	1	
	Happy	18	21	12	0	
	Sad	0	5	19	27	
III	Goat	1	0	0	18	19
	Pig	10	2	7	0	
	Cat	7	8	3	1	
	Mouse	1	9	9	0	
IV	Study	6	7	9	0	22
	Walk	0	3	3	16	
	Talk	13	3	4	2	
	Write	3	8	7	4	
V	Dirty	12	8	12	12	44
	Cold	23	15	5	1	
	Dark	4	13	14	13	
	Wet	5	8	12	19	

established by the procedure described in (2) above.

4. The five arrays of coefficients, yielded by the five replications, were treated according to conventional statistical procedures as indicated in Table II.

Table I above is included in this writing to show the degrees of agreement (disagreement) with regard to the rank-ordering of the words based upon estimated frequency of use.

An examination of Table II reveals that, in each replication, the proportion of positive coefficients is in excess of chance expectancy; and that in three of the five experiments, the difference between the obtained proportion and the chance expectancy proportion is significant well beyond the one percent level of confidence. Of the remaining two differences, one is shown to be significant beyond the five percent level. Perhaps it should be pointed out that this treatment yields conservative underestimates of levels of confidence, since it does not take into account the magnitudes of the rhos. The mean of the 175 coefficients was .45. Of these 175 coefficients, 143 were positive; two were zero and 30 were negative. Of the 143 positive coefficients, 16 attained unity (rho = 1.00). Of the 30 negative coefficients, only one attained unity.

INTERPRETATION AND CONCLUSIONS

The results reported in the preceding paragraphs clearly show that, as measured by the techniques described, the variety of the synonym repertoire varies directly with the frequency

TABLE II

COMPARISONS OF OBTAINED PROPORTIONS OF POSITIVE RHOS AND THE THEORETICAL PROPORTION OF POSITIVE RHOS TO BE EXPECTED IF THEY WERE NORMALLY DISTRIBUTED ABOUT ZERO AS A MEAN

Experiment	N	Obtained Proportion Positive	Theoretical Proportion Positive	Diff.	$\sigma_{Diff.}$	$\frac{Diff.}{\sigma_{Diff.}}$
I	39	89.74	50.00	39.74	9.36	4.24
II	50*	92.00	50.00	42.00	8.04	5.22
III	19	78.95	50.00	28.95	14.80	1.96
IV	21*	66.67	50.00	16.67	15.00	1.11
V	44	88.64	50.00	38.64	8.93	4.33
All	173	82.66	50.00	32.66	4.77	6.85

* One coefficient was zero.

with which a concept (word) is used. These findings are tentatively interpreted as confirming the hypothesis that the principle of reactive inhibition, as here extended, operates in the generation and use of synonyms. Other interpretations are cordially invited.

Summary

The series of five experiments, reported in the preceding sections, were designed to test the hypothesis that the principle of reactive inhibition, as amplified and extended in the foregoing pages, operates in the generation and use of synonyms. For the population of 175 students, used in these experiments, it was clearly shown that the variety of the synonym-vocabulary, as measured by the subject's report, varied directly with the frequency of word-usage as estimated by the subject. These findings are interpreted as confirming the hypothesis under investigation. They are consistent with the prediction that: when a concept (word) is used repetitiously by a person, or the persons of a culture, the synonym vocabulary representing the concept will be enriched; and that this enrichment will be proportional to the repetitiousness or frequency of use.

(Manuscript received for immediate publication February 23, 1950)

References

1. Hilgard, E. R. *Theories of learning.* New York: Appleton-Century, 1948.
2. Hull, C. L. *Principles of behavior.* New York: Appleton-Century, 1943.
3. Onions, C. T. (Ed.) *The Shorter Oxford English Dictionary.* Oxford: Clarendon Press, 1933.
4. Thorndike, E. L. *A teacher's word book of 20,000 words.* New York: Teachers College, 1932.
5. Thorndike, E. L. *The Thorndike-Century senior dictionary.* New York: Appleton-Century, 1941.
6. Thorndike, E. L., & Lorge, I. *The teacher's word book of 30,000 words.* New York: Teachers College, 1944.
7. Thorndike, E. L. On the frequency of semantic changes in modern English. *J. gen. Psychol.*, 1948, **39**, 23–27.
8. Zipf, G. K. The meaning-frequency relationship of words. *J. gen. Psychol.*, 1945, **33**, 251–256.
9. Zipf, G. K. *Human behavior and the principle of least effort.* Cambridge: Addison-Wesley, 1949.

THE RELATIONSHIP BETWEEN THE TILT OF A VISUAL FIELD AND THE DEVIATION OF BODY POSITION FROM THE VERTICAL IN THE WHITE RAT

BY BERNARD F. RIESS, HAROLD KRATKA AND ALBERT DINNERSTEIN

Animal Psychology Laboratory, Hunter College

Introduction.—It is easily observable that all higher animals have a sense of balance, that they can normally maintain an upright body position. In mammals there are four areas of sense data that contribute to this ability: visual, kinesthetic, labyrinthine, and tactile. It is reported that the kinesthetic and labyrinthine cues are primary and that vision plays a relatively unimportant role.

Part of the problem of the role of vision in body balance in animals is that, except in discrimination experiments, almost no work has been done to determine how animals see the world. In the discrimination situation, subjects are required to choose between a number of visual stimuli in order to receive food or escape punishment or both. Various devices have been used to test this perception-discrimination ability. Almost all the experiments that have been done have attempted to determine brightness, size, and form thresholds.

A limitation of discrimination experiments as a means of investigating vision in animals is that, while the method shows the relationship between the parts of a subject's visual field, it cannot demonstrate the role of visual field in relation to other sense modalities. Recent studies on the human organism, however, have attempted to correct this limitation. A series of papers by Witkin and Asch and by Witkin (1, 2, 3) on human subjects show the role of vision in the determination of body position. The apparatus consisted of a chair within a small room. Each could be tilted to the left or right independently by remote control from outside the room or by the subject seated in the chair. The experimenter adjusted the room and the subject to a given degree of tilt and the subject was instructed variously to adjust the room and the chair to the objective vertical. When the room was tilted and the subject was instructed to make his body vertical, most subjects would adjust the chair in the direction of the tilt of the room, showing that they used the orientation of the visual field as an important means of determining 'vertical' body position. The structure of the apparatus eliminated most of the kinesthetic cues because the subject, supported by the chair, was not required to use his muscles to keep himself from falling. Therefore, in this situation the task was to choose on the basis of either vision or labyrinthine, tactile and only weak kinesthetic factors.

Individual differences loomed large in Witkin's work. Some subjects relied mainly on visual, others mainly on their proprioceptive and tactile cues. It was found that women relied more on the visual field than did men. Asch and Witkin also modified previous experiments performed by Wertheimer and by Gibson and Mowrer. Here the subjects looked at a scene via a tilted mirror and were asked to adjust a rod in that scene to the true vertical. Asch and Witkin confirmed the results of the previous investigators, namely, that upon continual exposure to a conflict between the postural and visual cues to the vertical, the subject tends to see the tilted mirror scene in a subjectively upright position.

In another study Asch and Witkin seated the subjects in a tilted chair in front of a tilted room (2) and had them adjust a rod in that room to the vertical and horizontal. When the subjects viewed the room through a tube which concealed the surrounding laboratory,

upon continued viewing they at first increased and then slightly decreased their reliance on the visual field as represented by adjustment of the rod.

Purpose of present experiment.—In general, the experimental problem may be stated as the study of the role of a rotated visual field on the space-orienting behavior necessary for maintenance of upright body position by the white rat. One of the hypotheses examined here is that when labyrinthine, tactile and kinesthetic cues are not distorted, rats will tend to align themselves with the verticals indicated by their direct visual field. In other words, body position of the animal will deviate from the vertical as and in the direction which the visual field rotates. Furthermore, it is hypothesized that, as exposure to the conflict between visual and postural cues continues, the rats will increase their reliance on the visual mode of determining body position. It is our purpose to attempt to answer these hypothetical formulations and to determine whether there is a relationship in the degree of deviation from the vertical between the position of the animal and its visual field.

Subjects.—The subjects used in this experiment were 10 albino rats. Five males and two females of C-3 strain came from the laboratory colony at Carworth Farms. They were obtained at the age of 30 days and were used in this experiment at an approximate age of 120 days. The males had been used previously in learning experiments, one on a Warner-Warden alley maze and the other on a T-alley maze. Two females had previously been used on the Warner-Warden maze. The three remaining females came from the Sprague-Dawley colony of experimental rats and were obtained from the Rockefeller Institute at an approximate age of 150 days. These animals were used in this experiment about 20 days after receipt from the Institute, and had no behavioral study experience but had been used as a control group in dietary work. The 10 subjects had a mean weight of 213 grams with a range of 148–270 grams.

The subjects were housed in sexually segregated groups in wire cages, 10 × 20 × 10 in. Feeding was carried out routinely on a diet of 1½ standard Purina dog chow pellets daily. Water was supplied ad lib. Over the period of weekends, lettuce was added to the diet.

Apparatus.—In testing the present hypotheses, a procedure was required which isolated reliance upon kinesthetic factors from that upon visual cues in a situation demanding spatial orientation, a differentiation which could be recorded and measured. Drawing from Witkin's tilted room experiments (2), the deviation of the subjects from the vertical to the right or left was measured on a pivoted dowel-stick path under different conditions of rotation of a surrounding chamber which provided cues for visual orientation. Thus, the rats were made to run across a broomstick-like path leading through an enclosed chamber, which rotated independently of the central pathway. Fig. 1 describes the apparatus.

An enclosed rectangular plywood box, 36 × 12 × 12 in., had a hole 1¼ in. in diameter drilled through the center of each end piece. Just above each of these center holes, a No. 2½ tin can, opened at both ends, extended for ½ in. through an opening into the chamber. One of these cans was the entrance tube and the other an exit tube. Just adjacent to these and outside the main box were wooden starting and feeding platforms. The latter served as the goal box where food was obtained after each run and was enclosed on three sides by upright cardboard walls. The food consisted of whole-wheat bread crumbs in evaporated milk.

The center of the top cover of the chamber was a four in. wide dark tinted cellophane sheet which served to eliminate much of the outside illumination and allowed observation into the

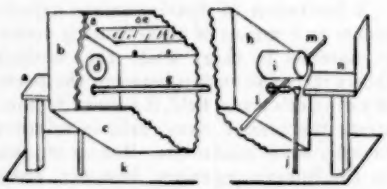

FIG. 1. Diagram of apparatus at 30° left tilt, with side cut away in order to show interior detail. a—starting platform; b—side of visual chamber; c—bottom of chamber; d—entrance tube; e—light bulb; f—tinted cellophane panel; g—dowel stick pathway; h—end piece (exterior); i—exit tube; j—pathway support; k—table; l—rod; m—stylus; n—feeding platform (with sides removed).

chamber. Inside the top of the chamber at the sides were eight flashlight bulbs, four on each side. These provided internal illumination. The bottom panel was placed on hinges so that a rat could easily be removed if it fell from the path, an occurrence which never happened. The box was suspended by wires from overhead braces allowing it to be rotated independently of the path bar.

The sides of the box were painted white and the top and bottom black. The end pieces were painted diagonally as if the walls extended on to each end, and so, with the top and bottom, forming four triangles with their apices meeting at the center holes in the ends. The triangles whose bases were continuous with the black bottom and top were painted black, and the triangles whose bases were continuous with the sides were white.

The path was a wooden dowel about 4 feet long and $\frac{3}{4}$ in. in diameter. This bar ran through the center holes in the end pieces, free from contact with them, and was suspended at each end in metal, V-cut, pivots fastened to upright supports. The bar was therefore free to rotate independently. A spring was stretched from the bar, at the entrance tube, to the table so that it provided a certain amount of tension in the bar and returned it to normal position after displacement. The approximate average displacement by the subjects throughout the experiment was 2° with a maximum of 7°. At the exit end of the box and under the exit tube, the protruding portion of the bar had a 12 in. metal rod inserted through its center at right angles to it. A stylus on this cross-bar registered all motion on a kymograph drum.

Procedure.—A. Training: To familiarize them with the path, the subjects were given 10 trials of which at least five had to be free of any pause midway on the bar. During the 10 trials, the bar was kept from rotating and the chamber was at 0° tilt. The subjects were divided into two groups, male and female, of five rats each.

B. Experimental: There were two directions of rotation of the field, to the right and to the left. In each of these directions, the following degrees of rotation from the vertical were established (with about 2° error) 0°, 10°, 20°, 30°. Each subject had four consecutive trials in each tilt situation. The order of degrees of rotation was randomly determined with the males and the identical order was then repeated with the female group. The first and last set of trials for all subjects were at 0° tilt.

For each trial, the subject was placed on the entrance platform near the entrance tube and allowed to enter. After having crossed the path, it ate for as long as it took to prepare a base line for the next recording on the kymograph, approximately 12 sec. Then the subject was carried back to the starting platform for the next trial until four trials were run. The experiment lasted for a period of 30 days.

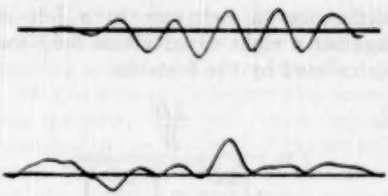

FIG. 2. Sample kymograph records. Upper graph shows trial at 0° tilt of apparatus. Lower graph shows trial at 10° right tilt of apparatus. Center line indicates true vertical. Curves show the animal's deviation from the vertical. Curve above center line shows right deviation. Upper, 0° trial, shows 49.5 percent right going tendency. Lower 10°, right trial, shows 80.7 percent right going tendency.

Results.—The graph of the animal's run is represented by an irregular line on the kymograph record (see Fig. 2). As the animal walked the bar, it shifted its weight to the right and to the left which is represented on the graph as that part of the wavy line which is above or below the base line respectively. When a subject's body position deviated consistently from the vertical, there is a corresponding greater part of the graph line above or below the base line.

The number of subjects did not warrant complex analysis of the graphic records with a planimeter. Therefore, the analysis is based on measurements of time as indicated by the length of the base line under the curves. The total length of the base line and the lengths of the base line under the curves indicating right deviation were measured. In order to compare trials of different absolute time, the time deviated right was divided by the total time, giving a percent right deviating tendency. The percentages were used as raw scores and the significance of the

difference in behavior in a left as against a right tilted visual field was calculated by the formula

$$t = \frac{\frac{\Sigma D}{N}}{\frac{\sqrt{\frac{\Sigma D^2}{N} - \left(\frac{\Sigma D}{N}\right)^2}}{\sqrt{N-1}}},$$

which is equivalent to $t = \frac{M_D}{\sigma M_D}$.

The results of the experimental trials are listed in the accompanying tables. The mean percent right deviating tendency is the mean of the percent time of each trial spent by the subjects in deviating towards the right. This shows in Fig. 1 as a curve above the base line. Thus, a mean percent greater than 50 percent indicates a right deviating tendency and a mean under 50 percent indicates a left deviating tendency. The figures in the second and third column in Table I indicate the mean per cent right deviating tendency at the various degrees of rotation of the visual field in left and right directions. At all degrees of rotation of the chamber, except for 10°, no significant difference was found between deviating tendency under conditions of right and left tilt, the level of significance ranging from values of P at .04 to P at .25. However, at 10° tilt, the difference in behavior in a right, as against a left tilted field was significant at .01.

TABLE I

Combined Means for All Subjects in Terms of Mean Percent of Time Spent Deviating to the Right from the Vertical and Differences Between Means for Left and Right Tilt for Four Conditions of Rotation

Degree of Box Rotation	Mean Percent Right Deviating Tendency		Mean Diff.	P
	Box tilted to right	Box tilted to left		
0°	54.1	51.8	2.3	.04
10°	58.8	41.7	17.1	.01
20°	51.2	48.5	2.7	.25
30°	54.1	45.7	8.4	.08

TABLE II

Mean Scores for Individual Subjects on Four Consecutive Trials in Terms of Mean Percent of Time Spent Deviating to the Right at 10° Tilt of Box

Subject	Mean Percent Right Deviating Tendency		Mean Diffs.	P
	Right Tilt	Left Tilt		
1	80.5	22.1	58.4	.02
2	62.1	31.4	30.7	.01
3	59.2	46.5	12.7	.20
4	55.9	49.7	6.2	.50
5	61.7	46.7	15.0	.07
6	60.2	40.2	20.0	.01
7	52.4	27.6	24.8	.04
8	41.9	46.8	4.8	.60
9	56.7	44.8	11.9	.63
10	58.2	54.3	3.9	.47

Table II gives individual scores for the 10 subjects and deals with the mean percent deviations and levels of significance of the differences between left and right rotation at 10° tilt. These differences are at a level of significance of .01 with respect to the trial of subjects No. 2 and No. 6 while the trial differences of subject No. 1 approaches significance with P at .02. Subject No. 7 registers a P of .04. The differences between trials of the other subjects give a much lower level of significance, with P ranging from .2 to .63.

In order to determine if experience within a given visual field increases or decreases a subject's reliance upon visual cues, the four trials at 10° tilt were grouped in order of occurrence (1st, 2nd, 3rd, 4th) and the mean percent deviations for all subjects on each trial, left and right, were plotted

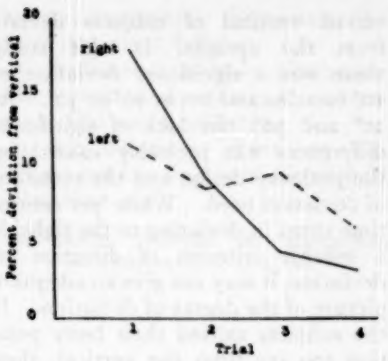

FIG. 3. Percent deviation from the absolute vertical according to trial, at 10° rotation of the apparatus; left and right tilt. Direction of the deviation of subjects is the same as the direction of tilt of the visual field.

on a graph. This results in two curves of learning within a group of trials and is shown in Fig. 3. To verify further the effects of learning in terms of significance, the first and fourth trials at 10° tilt were compared. The results are shown in Table III. When the field was rotated to the right, a level of significance of P at .01 was attained between the first and fourth trials. In the left rotated field, the level of significance was .2–.4. The significance of the difference between these trials at both left and right rotation combined reached a level with P at .01.

TABLE III

COMBINED MEAN SCORES FOR ALL SUBJECTS IN TERMS OF DEVIATION FROM THE VERTICAL ON THE FIRST AND LAST OF FOUR CONSECUTIVE TRIALS AT 10° ROTATION OF THE VISUAL FIELD

Condition of Visual Field	Mean Percent Right Deviating Tendency		Mean Diff.	P
	First Trial	Last Trial		
Box Tilted to Right	68.46	53.01	15.40	.01
Box Tilted to Left	38.14	43.83	5.73	.28
Total of Left and Right Tilts			10.54	.01

In order to see whether this learning persists, the 10° trials which occurred in the first half of the experiment were compared with 10° trials which occurred during the second half of the experiment. The significance of this difference reached a value of P at .1–.05. The significance of the difference between the sum of the mean differences of left and right tilt at 10° (see Table II) for comparison of males and females was at a level of .6–.4. Subjects 1–5 were males and subjects 6–10 were females.

Discussion.—The reliability of the difference in the direction of the deviation of the subjects from the vertical at 10° left and right rotation of the apparatus can be explained only in terms of the use of the verticals indicated by the visual field determinants of body position. Comparing the combined mean scores for 10 subjects in terms of the deviation from the vertical of the first and last of four consecutive trials at 10° rotation of apparatus, a significant decrease was shown in the effect of the visual field after repeated exposure to the conflict between visual and postural cues. This was probably caused by the strong postural cues which arose from the animals' walking in a non-vertical body position. In other words, the subjects learned to rely on their postural cues.

The results obtained seem to contradict the results of Wertheimer, Gibson and Mowrer, and Asch and Witkin (1), who all found that in the tilted mirror experiment, human subjects tended, with experience, to see the tilted mirror scene become vertical. In the tilted room experiment (2) there is some indication that the subjects, upon repeated viewing, in-

crease their reliance on the visual field while adjusting the rod to the vertical. The difference in findings between the present results and those of the previous investigators is not necessarily significant. In the mirror and tilted room experiment, what was being measured was the ability of the subjects to determine a visual vertical with the postural cues held constant. In this study the problem was one of establishing a postural vertical within a tilted visual field. The former is a question of the usefulness of postural cues in the determination of a visual vertical, the latter is a question of the effect of visual cues on the postural vertical. There is not necessarily a direct relationship between the two phenomena.

If Witkin should compute the relationship between successive trials in the experiment in which he had the subjects adjust the chair in a tilted room to the vertical, it could be established if this difference in results arises from the species of the subject or from the experimental design.

When the subjects in this present study were compared according to sex, it was found that at 10° rotation the females showed less deviation than did the males, but this difference was not significant. This may indicate a lack of sex difference in perception in rats, but the results could also be attributed to the small sample or the dissimilar histories of the two groups of animals. It should be noted that the Witkin study did not establish whether the observed difference was primarily biogenic or sociogenic.

In the experiments by Asch and Witkin and by Witkin, it was found that the greater the rotation of the visual field the more did the perceived vertical of subjects deviate from the upright. In this study there was a significant deviation at 10° rotation and not at 20° or 30°. At 20° and 30°, the lack of significant differences was probably caused by the pathway design and the standard of deviation used. While 'per cent of time spent in deviating to the right' is a reliable criterion of direction of deviation, it may not give an adequate picture of the degree of deviation. If the subjects shifted their body position too far from the vertical, they would start to slide off the pathway. The very strong postural cues thus aroused cause them to correct their deviation. At 20° and 30° there must have been postural cues which did not occur at 10°. The mean difference between left and right tilt is largest at 10°, decreases sharply at 20°, and increases slightly at 30°. While it is probable that the direct relationship between the tilt of the visual field and the degree of shift of the perceived upright found by Witkin in humans, also exists in rats, this could not be shown by this apparatus. Witkin measured deviation directly in degrees and his subjects were securely supported by the chair. To get similar results with rats would require a technique to measure *intensity* of deviation and a pathway which prevented objective loss of balance.

Summary

1. Rats were subjected to a tilted visual field to which they had to respond by locomotor adjustment.

2. Reliable differences in deviation from the vertical were obtained when the visual field was displaced 10° from the horizontal in the right and left direction. The direction of devia-

tion was consistently the same as the tilt of the visual field.

3. A decrease in body dislocation with increased experience in a tilted visual field was established.

4. Sex differences were not significant for any degree of visual field dislocation.

(Manuscript received for immediate publication June 5, 1950)

References

1. Asch, S. E., & Witkin, H. A. Studies in space orientation: I. Perception of the upright with displaced visual fields. *J. exp. Psychol.*, 1948, **38**, 325–337.
2. Asch, S. E., & Witkin, H. A. Studies in space orientation: II. Perception of the upright with displaced visual fields and with body tilted. *J. exp. Psychol.*, 1948, **38**, 455–477.
3. Witkin, H. A. Sex differences in perception. *Trans. N. Y. Acad. Sci.*, 1949, **12**, 22–26.

PSYCHOLOGICAL MONOGRAPHS

#270

1945

ON PROBLEM SOLVING

By

KARL DUNCKER

$2.50

This popular monograph is #270 of the Psychological Monograph series. It has been reprinted so that it is again available.

AMERICAN PSYCHOLOGICAL ASSOCIATION

1515 Massachusetts Avenue N. W.

Washington 5, D, C.

RULES REGARDING ANIMALS

a statement prepared by

The Committee on Precautions
in Animal Experimentation
of the
American Psychological Association

and approved by

The Council of Representatives of the
American Psychological Association

on September 8, 1949

COPIES SUITABLE FOR FRAMING

sent to

Animal Laboratories at Their Request